Automatic Trading using Stochastic Methods

by

Hui Gong MSc

A thesis submitted in conformity with the requirements

for the degree of Doctor of Philosophy

Department of Mathematics

Faculty of Mathematical & Physical Sciences

University College London

December, 2018

Disclaimer

I, Hui Gong, confirm that the work presented in this thesis is my own. Where information and data have been derived from other sources, I confirm that these have been indicated in the thesis.

Signature _____

Date _____

Abstract

In this thesis, we develop algorithms for automatic trading and execution strategies for institutional investors. In the first part, we develop optimal execution strategies for traders who trade continuously using only market orders and account for stochastic trading impact. There are a great variety of impacts in the electronic trading market which may affect the performance of trading strategies in a direct or indirect manner. To understand the way of measuring and taking control of the effects potentially caused by these impacts, some of traders opt to simulate the impacts by using mathematical models such as stochastic control theories. These attempts help traders to find solutions, such as how to develop an optimal execution strategy by solving Hamilton-Jacobi-Bellman equations and how these strategies affect trading. In the second part, we focus on a new market, the cryptocurrencies' market, and find out the pairs trading strategies for the buy-side investors.

We introduce the traditional trading model, Almgren-Chriss model in Chapter 2, and use it to benchmark the performance of the strategies we proposed. Chapters 3 and 4 illustrate how agents or sell-side traders interact in the market when stochastic market impacts and latency impact are modelled. We also explore the numerical methods and closed-form expression to obtain the optimal execution strategy. In Chapter 5, we analyse how to execute by using co-integrated pairs trading as a buy-

side trader in the cryptocurrencies' market. We consider how to trade 'BTC/USD' and 'ETH/USD' by using the quantitative trading methods and find out the optimal weight for each cryptocurrency.

This thesis was completed under the supervision of ***Professor Álvaro Cartea****.*

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor Professor Álvaro Cartea for his kind guidance and supervision during the past five years. His insights and experience encourage me to constantly improve myself.

Secondly, I would like to thank Professor Sebastian Jaimungal, University of Toronto for his brilliant academic help and technical support. His ideas and advice deepen my understanding of this area.

Thirdly, I would like to thank all the supporting staff at the Department of Mathematics, UCL for their patience when I had inquires regarding administrative matters.

Finally, I would like to thank my parents for their continuous support.

Hui Gong, *University College London*, December 2018

To my beloved parents Jin Gong, Zhiming Gao and my wife Sha Sha.

It has been my experience that competency in mathematics, both in numerical manipulations and in understanding its conceptual foundations, enhances a person's ability to handle the more ambiguous and qualitative relationships that dominate our day-to-day financial decision-making.

— Alan Greenspan

Innovation must lead infrastructure for a simple but compelling reason: Innovation produces new types of products and markets, and it is virtually impossible to know how to run those markets efficiently before they are created.

— Myron Scholes

Quantitative Finance is very much an empirical field. You can have the world's best theoretical model, but if you cannot justify it from real data, it isn't going to be much use.

— Robert Almgren

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

With the enormous development in computing technology, it is both necessary and unavoidable for the financial industry to apply mathematical models and computer science technologies to pricing financial products, trading underlying assets and managing risks. For example, computer science technology has already been employed in NASDAQ electronic exchange to execute orders. Due to its simplicity and low requirement of computer equipment, time-weighted average price (TWAP) and volume-weighted average price (VWAP) strategies can no longer satisfy the demand of the agents who expect more complicated models to implement.

Also, market participants are required to implement stress tests in line with BASEL III and MiFID II. Therefore, to better understand how the black box of algorithmic trading and the trading impacts work, it is necessary for us to define and differentiate quantitative trading, algorithmic trading (AT) and high frequency trading (HFT).

Here, I define the trading as three types:

- *Quantitative trading*: Using the mathematical or statistical models to price the derivatives and decide the hedging ratio. It is normally used to deal with the

problem in a **static** form.

- *Algorithmic trading*: Buying or selling the underlying assets via the algorithm solved by **dynamic** methods. We set up an objective function to measure the trading performance and decide execution direction (long/short), time, frequency and volume. Regarding frequency, it can either be low or high.

- *High-frequency trading*: Trading in high-frequency to seek opportunities based on the algorithmic trading theories. These trading opportunities cannot be realised by simply buying low and selling high. Actually it requires the participants to make a **quick response to the market change**. In that case, the agents will post orders to the limit order books (LOBs) more frequently than their counterparts or post several marker orders (MOs) in a very short time interval.

There are two fundamental types of orders in an order driven electronic market: *Market orders* (MOs) and *Limit Orders* (LOs). A market order is an order to acquire or liquidate shares immediately at the best price available at the moment when the order is received by the trading system. It is widely used by agents whose primary concern is to get the trade done immediately regardless of the price to be executed at. In short, a market order guarantees the order's immediate execution instead of a particular price. For those agents who intend to capture a specific price or a better price observed from the order book, they may choose another type of order called 'limit orders'. By using a limit order, agents can set maximum or minimum price at which they are willing to buy or sell shares, which means the orders can only be executed when the price satisfies the conditions set beforehand by the agents. In this regard, a limit order is unlikely to be executed if the price fails to touch the threshold throughout the trading period. To put it in a nutshell, a limit order, contrary to a market order, is a passive type of order that can guarantee a price instead of the

execution.

In this thesis, we assume that the agents use only **MOs** to do the execution and discuss the price impact as a result of executing MOs. If we ignore price impact, it is likely that it can significantly influence the performance of the algorithmic trading strategies. Specifically, market impact refers to the immediate changes in price and the change in near future as a reaction of the market to an incoming order. In other words, it describes the 'causality' between the incoming order and the subsequent price change. Here, we consider two price impacts: temporary price impact and permanent price impact. The temporary price impact is caused by market depth, which refers to the available volume posted at different levels of the LOBs. The permanent price impact is related to the downward (upward) pressure on the price of the stock when agents liquidate (acquire) a large volume of shares.

Apart from the price impacts mentioned above, the agents have to face latency impact as a result of different physical venues and computer devices. In this thesis, we use the stochastic differential equations to model the impacts and determine the corresponding optimal execution strategies.

This thesis looks at the pairs trading strategies in cryptocurrencies' market, a new market which tends to disrupt the existing ways of fund-raising and the secondary markets.

1.2 Literature Review

Here is a brief review of studies on general algorithmic trading and optimal executions. The detailed reviews are followed in the specific chapters, including those of the price impacts, latency impact and pairs trading.

Our study on optimal execution strategies starts with the basic elements of electronic markets and the main ways through which people participate in and interact

with the market. Technological improvements enable average agents to arrive at informed decisions based on their trading ideas by making use of state-of-the-art trading algorithms and quantitative analysis. The trading algorithms designed by mathematical tools can accurately map a diachronical market behavior and be used by agents as an empirical indicator of future results while making decisions [28]. Such change in the ways of trading is largely driven by technology evolution which has enabled algorithmic and high-frequency trading and the modernization of financial market. Hence, more attention has been given to the studies on market microstructure and market impact on the price formation process and optimal trading strategies [20]. Traders can also enjoy considerable benefits delivered by the algorithmic execution as it has become an inseparable part of today's financial market, including the lower volatility and execution costs, as well as higher market stability and transparency [4]. In terms of execution, despite the explosion of various derivative orders types, two core order types are still most frequently used in today's electronic market: limit order and market order [1]. The orders are managed by a matching engine and a limit order book (LOB) following price-time priority, also known as the first-in-first-out (FIFO) execution schedule. LOB keeps a record of incoming and outgoing orders received by the trading system, while the matching engine triggers their execution by using predefined algorithms. The codes and instructions are embedded in the algorithms. Each code or instruction targets a specific goal, such as establishing when a possible execution can take place and what conditions need to be satisfied when selecting the orders which will be executed [4]. Although there are a wide variety of trading algorithms which are evolving in a similar fashion to electronic trading, they can be generally categorized into three main types, namely impact-driven algorithms, cost-driven algorithms and opportunistic algorithms, with the aims to minimize the overall market impact, to reduce the overall trading costs and to take advantage of the favorable market conditions respectively [19]. The most

common impact algorithms are named time weighted average price (TWAP) and volume weighted average price (VWAP). The cost-driven algorithms do not take into account market impact, timing risk, and other related factors such as the measures of benchmarking and implementation shortfall. The design of a proper opportunistic algorithm should take into account the following: (1) the fee structure associated to each trading venue; (2) the latency, which in this case represents the time lag between orders sent and processed from each single venue; and (3) the probability of execution associated to each trading venue [19].

The algorithms designed for optimal execution in the electronic market are based on numerous sophisticated mathematical analysis and modelling tools. To find out optimal strategies with mean-reverting price, Leung and Li delve into three important mean-reverting models: Ornstein-Uhlenbeck (OU), exponential Ornstein-Uhlenbeck (XOU), and Cox-Ingersoll-Ross (CIR) models [22]. Because of their tractability and interpretability, all these models are widely used in describing and estimating mean reversion in asset prices. Apart from the mean-reverting processes, the modern probability theory and ergodic theory of stationary stochastic processes are also used in HFT by agents believing in technical analysis [32]. Mean reversion and stationarity are considered as two equivalent ways of examining the same type of price series. Although only few price series are found to be mean reverting, we can create more by combining two or more individual price series that are not mean reverting into a portfolio whose price turns out to be mean reverting [11]. ØKsendal and Sulem provides Hamilton-Jacobi-Bellman (HJB) Equation with a solution through dynamic programming method, which is the most important and useful solution method of stochastic control problems that can be set about when developing optimal execution strategies [27]. We mainly use their method to figure out the problems we propose in this thesis.

A lot of scholars have also studied optimal execution strategies in the context

of algorithm trading. The 'optimal execution problem' is encountered by investors who seek to execute large orders over a given trading horizon and whose actions give rise to impacts on the market price. The problem is mainly concerned with how to develop a trading strategy that maximize an appropriate objective function, and its key issue is to model the price impact of share trading [16]. Among these, the breakthrough outcome of Almgren and Chriss's study was regarded as the first example to directly address the issue of permanent and temporary market impact in a continuous time setting, and the thesis is considered as one of the milestones within the literature spectrum of high frequency optimal execution [2]. Since then, this area have attracted considerable attention, and numerous extensions have been added to the Almgren-Chriss execution strategy. Álvaro Cartea incorporated the permanent impact of market order-flow in an close-form execution strategy, which consists of an Almgren-Chriss execution strategy adjusted by a weighted-average of the future expected net order-flow [7].

1.3 Main Results & Outline

This thesis mainly focuses on the mathematical aspects of algorithmic trading. It establishes mathematical frameworks for optimal execution strategies. It serves as an specification for agents when they encounter various types of impacts during their trading. The main theoretical tools employed are mean reverting process, numerical analysis and stochastic optimal control theory.

In Chapter 2, the Almgren-Chriss (AC) execution strategy is introduced, which we employ as a benchmark.

Chapter 3 discusses the model for price impacts and how the agents balance price risk and price impacts. When the agents execute a large number of shares via MOs, they will leave sustained price impacts. MOs produce two types of impacts on mid-

prices. Temporary impact refers to the order walking through the different levels of the limit order books (LOBs), and after a period of time (milliseconds) the LOBs can be replenished, depending of the elasticity of the price-formation process due to liquidity (average, instantaneous, hidden). Permanent impact refers to the enduring changes in the mid-prices due to the information conveyed by the MOs which are impounded in the price of the asset. Chapter 3 assumes that they are modelled as a stochastic process. We also analyse the optimal execution strategies when price impacts are stochastic, then compare the results to the benchmark in Chapter 2. We estimate model parameters by using historical NASDAQ millisecond-stamped messages and show the performance of strategies developed here.

In Chapter 4, we define latency as the delay time between a signal and a response because of the time it takes for the signal to travel inside the automated trading system. This latency concerns traders since the quoted price of the asset may fluctuate within this period, making the price initially captured by the trader the moment he/she triggered the order different from the one when the order was eventually received and executed by the exchange. The difference between the observed price and execution price is referred to as latency impact. Latency impact is a stochastic value affecting the trader's trading behaviours. We provide an explicit closed-form strategy for traders who seek to optimally liquidate or acquire shares with the stochastic impact of latency taken into account. We assume the latency to be an Ornstein-Uhlenbeck (OU) type process with jumps and employ an OU process to formulate latency impact as it follows an auto-regression model in the discrete form. The optimal trading speed in our strategy is generated by a dynamic programming problem and is found to be affected by latency impact in a linear form. We use historical data to calibrate the parameters in our strategy and we compare its performance with that of the benchmark in Chapter 2.

In Chapter 5, we develop a pairing price model for the existed mainstream cryp-

tocurrencies, including Bitcoin (BTC) and Ethereum (ETH). We introduce a short-term alpha, which is the co-integrated factor between the pairs of 'BTC/USD' and 'ETH/USD'. Base on the this stochastic factor, we set the objective function is exponential utility of the agent's terminal wealth. Finally, we find an optimal closed form amount for each underlying asset in this pair and compare the results based on the simulations of different levels of the risk aversion.

The models in Chapter 3 and Chapter 4 derive from the models designed by Álvaro Cartea in his book [4]. I make improvements on this basis. The model in Chapter 5 was independently proposed by me. What Álvaro Cartea mentioned in [8] is a high-order model, and I applied the most basic situation of this model to cryptocurrencies.

Chapter 2

Methodology and Benchmark

2.1 Stochastic Control Problem

This section focuses on methodology and the problems we are trying to solve. With reference to Cartea's book [4], stochastic control problems are reflected and applied in many aspects of financial modeling, for example, the optimal investment in continuous time proposed and solved by Merton [24], which is regarded as one of the classic applications of the stochastic control problems. There are also many other applications, such as optimal dividend setting, optimal entry and exit problems, utility indifference evaluation and so on. Despite the broad applications in different aspects, the core goal of the stochastic control problems is to maximize (or minimize) certain expected profit (cost) functions by adjusting their own strategies that influence the dynamics of the underlying stochastic system. Find the strategy to reach the maximum (or minimum).

The agent's wealth is affected by her behaviour on one hand and modulated by the uncertain dynamics in the traded assets in a stochastic manner on the other hand, while the resulting optimal strategies tied to the dynamics of the assets may also give effects on the wealth. It is worth noting that, the optimal strategies turn out to become Markov in the underlying state variables in many cases, even though

the agent has already taken non-Markovian controls into consideration which may depend on the entire history of the system.

The **dynamic programming principle** (DPP) and the related non-linear partial differential equation (PDE) known as the **Hamilton-Jacobi-Bellman** (HJB) equation - also called the **dynamic programming equation** (DPE), is a key tool frequently used to solve the stochastic control problems.

2.1.1 The Optimal Liquidation Problem

We imagine a scenario where the agent holds a large amount of shares $\mathfrak{N}$ of an asset at the price of S_t, and it is no longer a value investment for her to hold based on her fundamental analysis on the assets. As a result, the agent seeks to liquidate the shares by the end of day (defined as time T in this case). However, the agent is high likely to obtain poor prices if she tries to liquidate all the shares immediately since market cannot provide ample liquidity to absorb a large sell order at the best available price. In this case, a usual action that an agent may take to address the issue is to spread this out over time and solve a stochastic control problem. Also, the agent may be urgent to get rid of these shares by penalizing holding inventories different from zero in the whole strategy. it ν_t denotes the rate at which the agent sells her shares at time t, then the agent seeks the value function

$$H(x,S,q)^{\nu} = \sup_{\nu \in \mathcal{A}_{\prime,\mathcal{T}}} \mathbb{E}\left[\underbrace{X_T^{\nu}}_{\text{Ternimal Cash}} + \underbrace{Q_T^{\nu}\,(S_T^{\nu} - \alpha\, Q_T^{\nu})}_{\text{Terminal Execution}} - \underbrace{\phi \int_t^T (Q_s^{\nu})^2 ds}_{\text{Inventory Penalty}} \right] , \tag{2.1}$$

and the resulting optimal liquidation trading strategy ν^*, where,

$$dQ_t^{\nu} = -\nu_t\, dt\ , \qquad q_0^{\nu} = \mathfrak{N}\ , \qquad \text{agent's inventory,} \tag{2.2}$$

$$dS_t^{\nu} = -g(\nu_t)\, dt + \sigma dW_t\ , \qquad S_0^{\nu} = S\ , \qquad \text{fundamental asset price,} \tag{2.3}$$

$$\widehat{S}_t^\nu = S_t^\nu - f(\nu_t) \qquad \widehat{S}_0^\nu = S \ , \qquad \text{execution price,} \tag{2.4}$$

$$dX_t^\nu = \nu_t \, \widehat{S}_t^\nu \, dt \qquad X_0^\nu = x \ , \qquad \text{agent's cash.} \tag{2.5}$$

In the above,

- $\nu = (\nu_t)_{0 \le t \le T}$ is the trading speed, which is the variable that agent controls to liquidate or acquire shares in the optimization problem and ν_t^* denotes the optimal rate,
- $Q^\nu = (Q_t^\nu)_{0 \le t \le T}$ is the agent's inventory, which is affected by how fast the agent trades,
- $W = (W_t)_{0 \le t \le T}$ is a Brownian motion,
- $S^\nu = (S_t^\nu)_{0 \le t \le T}$ is bid-price process, and is affected primarily by the trading rate as well,
- $\widehat{S}^\nu = (\widehat{S}_t^\nu)_{0 \le t \le T}$ is the execution price process, which the agent can sell by walking the LOB,
- $X^\nu = (X_t^\nu)_{0 \le t \le T}$ is cash process resulting from the agent's execution strategy.
- g, $f : \mathbb{R} \to \mathbb{R}^+$ denote the **permanent** and **temporary** (negative) price impact functions that agent's trading action has on the fundamental price and execution price respectively,
- $\mathcal{A}_{t,T}$ is the admissible set of strategies: $\mathcal{F}$- predictable non-negative bounded strategies.

2.1.2 Dynamic Programming Equation / Hamilton-Jacobi-Bellman Equation

As the agent always seeks to maximise the value function, we define the **performance criteria** as follows,

$$H(t,x) := \sup_{u \in \mathcal{A}_{t,T}} H^u(t,x) \ , and$$
$$H^u(t,x) := \mathbb{E}_{t,x}\left[G(X_T^u) + \int_t^T F(s, X_s^u, u_s) ds \right] \ ,$$

where the notion $\mathbb{E}_{t,x}[\cdot]$ represents expectation conditional on $X_t^u = x$. These two objects are the time indexed analog of the original control problem and the performance criteria.

Theorem 2.1.1 ***Dynamic Programming Principle (DPP).*** *The value function 2.1.2 satisfies the DPP*

$$\boxed{H(t,x) = \sup_{u \in \mathcal{A}_{t,T}} \mathbb{E}_{t,x}\left[H(\tau, X_\tau^u) + \int_t^\tau F(s, X_s^u, u_s) ds \right]}, \tag{2.6}$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ *and all stopping times* $\tau \leq T$.

This equation is really a sequence of equations that tie the value function to its future expected value, plus the running reward/penalty. And the DPE is an infinitesimal version of the DPP. There are two key ideas involved:

- Setting the stopping time τ in the DPP to be the minimum between (a) the time it takes the process X_t^u to exit a ball of size ϵ around its starting point, and (b) a fixed (small) time h - all while keeping it bounded by T.

- Writing the value function (for an arbitrary admissible control u) at the stopping time τ in terms of the value function at t using Itô's lemma. Specifically,

assuming enough regularity of the value function, we can write

$$H(\tau, X_\tau^u) = H(t,x) + \int_t^\tau (\partial_t + \mathcal{L}_s^u) H(s, X_s^u) ds + \int_t^\tau \mathcal{D}_x H(s, X_s^u)' \sigma_s^u dW_s \ , \quad (2.7)$$

where $\sigma_t^u := \sigma(t, X_t^u, u_t)$ for compactness, $\mathcal{L}_t^u$ represents the infinitesimal generator of X_t^u, and $\mathcal{D}_x H(\cdot)$ denotes the vector of partial derivatives with components $[\mathcal{D}_x H(\cdot)]_i = \partial_{x^i} H(\cdot)$. For example, in the one-dimensional case,

$$\begin{aligned} \mathcal{L}_t^u &= \mu_t^u \partial_x + \frac{1}{2}(\sigma_t^u)^2 \partial_{xx} \\ &= \mu(t,x,u)\partial_x + \frac{1}{2}\sigma^2(t,x,u)\partial_{xx} \end{aligned}$$

Suppose that $u^*$is an optimal control, then from Equation 2.6, we have

$$H(t,x) = \sup_{u \in \mathcal{A}_{t,T}} \mathbb{E}_{t,x} \left[H(\tau, X_\tau^{u^*}) + \int_t^\tau F(s, X_s^{u^*}, {u^*}_s) ds \right] \ . \quad (2.8)$$

By applying Itô's lemma to write $H(\tau, X_{\tau)}^{u^*})$ in term of $H(t,x)$ plus the integral of its increments, taking expectations, and then the limit as $h \searrow 0$, we find that

$$\partial_t H(t,x) + \mathcal{L}_t^{u*} H(t,x) + F(t,x,u^*) = 0 \ . \quad (2.9)$$

We finally arrive at the DPE (also known in this context as the **Hamilton-Jacobi-Bellman**)

$$\boxed{\begin{aligned} \partial_t H(t,x) + \sup_{u \in \mathcal{A}} (\mathcal{L}_t^u H(t,x) + F(t,x,u)) = 0 \ , \\ H(T,x) = G(x) \ . \end{aligned}}$$

The terminal condition above follows from the definition of the value function in Equation 2.1.2 from which we see that the running reward/penalty drops out an $G(X_T^u)$ is $\mathcal{F}_T$ measurable.

2.1.3 Riccati Equation

In most cases, we need to solve the Riccati equation, named after Jacopo Riccati, to find out the solution of the DPE. A Riccati equation is any first-order ordinary differential equation that is quadratic in the unknown function, i.e. an equation of the form

$$h'(t) = q_0(t) + q_1(t)\, h(t) + q_2(t)\, h^2(t) \ , \tag{2.10}$$

where h is a function of t and h' is the first order derivative. The function $q_0(t) \neq 0$ and $q_2(t) \neq 0$. If $q_0(t) = 0$ the equation reduces to a Bernoulli equation, while if $q_2(t) = 0$ the equation becomes a first order linear ordinary differential equation. Normally, we will have a terminal condition for the value function at time T, i.e. one particular solution h_* can be found. Then the general solution is obtained as $h(t) = h_* + u(t)$. Substituting this into the Riccati equation, we have

$$h'_* + u' = q_0 + q_1 \cdot (h_* + u) + q_2 \cdot (h_* + u)^2 \ . \tag{2.11}$$

Since

$$h'_* = q_0 + q_1\, h_* + q_1\, h_*^2 \ , \tag{2.12}$$

$$u' = q_1\, u + 2q_2\, h_*\, u + q_2\, u^2 \ , \tag{2.13}$$

i.e. there is only one Bernoulli equation left, which is

$$u' = -(q_1 + 2q_2\, h_*)\, u = q_2\, u^2 \ . \tag{2.14}$$

To solve this equation, we set a new substitution $v(t) = \frac{1}{u(t)}$, then Equation 2.14 changes to

$$v' = -(q_1 + 2q_2\, h_*)\, v = -q_2 \ . \tag{2.15}$$

A set of solutions to the Riccati equation is then given by

$$h(t) = h_* + \frac{1}{v(t)} \,, \tag{2.16}$$

where v is the general solution to the aforementioned linear Equation 2.15.

2.2 Almgren-Chriss Model

In this section, the classical Almegren-Chriss execution model is introduced as a benchmark of our model. The agent (trader) only uses MOs to liquidate (acquire) a large quantity of $\mathfrak{N} > 0$ shares, but he/she will be faced with both a temporary price impact and a permanent price impact. As mentioned before, both impacts are deterministic and assumed to be linear with respect to the quantity traded.

Supposing $Q_0 = \mathfrak{N}$ amount of shares is liquidated over the time period $[0, T]$, and MOs are send at speed ν_t, then the inventory denoted by Q_t^ν, follows

$$dQ_t^\nu = \pm\, \nu_t \, dt \,, \qquad\qquad Q_0^\nu = \mathfrak{N} \,. \tag{2.17}$$

The agent controls the rate ν.

In the equation 2.17, the sign $\pm$ depends on whether the problem is liquidating $(-)$ or acquiring $(+)$ shares.

The trading process will be exposed to the temporary and permanent price impact as follows.

2.2.1 Optimal Execution with Constant Price Impact

Here, we only focus on the problem of liquidating the asset, while acquiring the asset is similar. The bid price (ask price in the case of acquisition) satisfies the SDE

with a constant permanent price impact

$$dS_t^\nu = -b^C \, \nu_t \, dt + \sigma \, dW_t \ , \tag{2.18}$$

where S_t^ν is the bid price process, which is affected by the trading rate ν_t. b^C is the constant, linear permanent impact that the agent's trading action made on the price, σ is the volatility component and W_t is a standard Brownian motion.

The execution price $\widehat{S_t}$ also has a constant, linear temporary impact $k^C \geq 0$, i.e.,

$$\widehat{S_t} = S_t^\nu - k^C \, \nu_t \ , \tag{2.19}$$

where k^C denotes the temporary price impact that the agent's trading action made on the price they can execute the trade at.

The agent's cash process X_t^ν satisfies the SDE

$$dX_t^\nu = (S_t^\nu - k^C \, \nu_t) \, \nu_t \, dt \tag{2.20}$$

We assume that the agent's performance criterion is given by

$$H^\nu(t,x,S,q) = \mathbb{E}_{t,x,S,q}\left[\underbrace{X_T}_{Ternimal\ Cash} + \underbrace{Q_T \, (S_T - \alpha \, Q_T)}_{Terminal\ Execution} - \underbrace{\phi \int_t^T (Q_u^\nu)^2 du}_{Inventory\ Penalty} \right] \ ,$$

and the expectation is conditional on $X_t = x$, $S_t^\nu = S$, and $Q_t = q$. The terminal execution penalty coefficient is $\alpha \geq 0$ and the inventory penalty coefficient is $\phi \geq 0$.

The agent's value function is

$$H(t,x,S,q) = \sup_{\nu \in \mathcal{A}} H^\nu(t,x,S,q) \ ,$$

where $\mathcal{A}$ is the set of admissible strategies: $\mathcal{F}$-predictable non-negative bounded

strategies.

2.2.2 Hamilton-Jacobi-Bellman (HJB) equation and its solution

Applying the dynamic programming principle (DPP), the value function satisfies the dynamic programming equation (DPE):

$$0 = \partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H - \phi\, q^2 + \sup_{\nu} \left[(S - k^C\, \nu)\, \nu \partial_x H - b^C\, \nu\, \partial_S H - \nu\, \partial_q H \right] \ , \quad (2.21)$$

subject to the terminal condition

$$H(T, x, S, q) = X_T + q_T\, S_T - \alpha\, q_T^2 \ . \quad (2.22)$$

The solution has been originally proposed by Almgren and Chriss. Here, I confirm it by providing a detailed derivation process.

Proposition 2.2.1 *The DPE (2.21) admits the solution*

$$H(t, x, S, q) = x + q\, S + h(t)\, q^2 \ , \quad (2.23)$$

where

$$h(t) = \sqrt{k^C\, \phi} \frac{1 + \zeta\, e^{2\,\gamma (T-t)}}{1 - \zeta\, e^{2\,\gamma (T-t)}} - \frac{1}{2}\, b^C \ ,$$

with $\gamma = \sqrt{\frac{\phi}{k^C}}$ *and* $\zeta = \frac{\alpha - \frac{1}{2}\, b^C + \sqrt{k^C\, \phi}}{\alpha - \frac{1}{2}\, b^C - \sqrt{k^C\, \phi}}$.

Proof. To solve the DPE (2.21), we need to find the initial optimal trading speed first

$$\nu_* = \frac{1}{2\, k^C} \frac{(S \partial_x - b^C \partial_S - \partial_q)\, H}{\partial_x\, H} \ . \quad (2.24)$$

Then by substituting the ansatz into (2.21), we get the non-linear PDE

$$0 = (\partial_t h - \phi)\, q^2 + \frac{1}{4k}\left[b\, q + 2\, q\, h(t)\right]^2 \ . \tag{2.25}$$

Dividing q^2, we can get the coefficient who satisfies the non-linear ODE

$$0 = \partial_t h - \phi + \frac{1}{k}\left[h(t) + \frac{1}{2}\, b\right]^2 \ , \tag{2.26}$$

which is a Riccati type and can be integrated directly. Let $h(t) = \chi(t) - \frac{1}{2}\, b$, we obtain

$$\frac{\partial_t \chi}{k\, \phi - \chi^2} = \frac{1}{k} \ , \tag{2.27}$$

subject to $\chi(T) = -\alpha + \frac{1}{b^C}$ Then after integrating both sides of the above from t to T, we will get the solution

$$\chi(t) = \sqrt{k^C\, \phi}\frac{1 + \zeta\, e^{2\,\gamma(T-t)}}{1 - \zeta\, e^{2\,\gamma(T-t)}} \ , \tag{2.28}$$

where $\gamma = \sqrt{\frac{\phi}{k^C}}$ and $\zeta = \frac{\alpha - \frac{1}{2}\, b^C + \sqrt{k^C\, \phi}}{\alpha - \frac{1}{2}\, b^C - \sqrt{k^C\, \phi}}$. □

Theorem 2.2.2 ***Verification.*** *The function provided in Proposition 2.2.1 is the classic solution of the DPE in equation (2.21). And the trading rate is given by*

$$\nu_t^* = \gamma\, \frac{\zeta\, e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta\, e^{\gamma(T-t)} - e^{-\gamma(T-t)}}\, q_t^{\nu^*} \ , \tag{2.29}$$

where γ and ζ is defined in Proposition 2.2.1.

Proof. Function (2.23) is the first and second order differentiable to its variable t, x, S and q. It also satisfies the DPE (2.21). We also verify the optimal strategy ν^* constructed from the admissible solution set, i.e. $\int_0^T \mid \nu^* \mid dt < \infty$, which means the strategy is obviously integrable. □

Chapter 3

Stochastic Price Impacts

3.1 Introduction

As we mentioned in Chapter 1, in electronic trading markets, agents may often find themselves in a situation where they need to buy or sell a great quantity of shares, more than the current available liquidity in the LOB. In such scenario, the parent order needs to be sliced into smaller, child orders, and the trader needs to assess the effect of his early orders on the later orders. The amount of shares we are referring to is too large to execute in one trade [4].

Thus 'optimal execution problem' in this chapter is an issue encountered by investors who seek to execute a large order over a given trading horizon and whose actions give rise to impacts on the market price. If the agent executes large orders, he/she will bear direct and indirect cost. The indirect cost includes the price impact, which is quite difficult to quantify. In this chapter, we define the price impact from two perspectives, namely temporary impact and permanent impact. Temporary impact refers to the order walking through the different levels of the limit order books (LOBs). The LOBs may be replenished after a period of time (milliseconds), otherwise it contributes to the permanent impact.

Permanent impact refers to changes in the prices that are generated over a pe-

riod of five minutes as a result of the information conveyed by the Market Orders (MOs) which are impounded in the price of the asset [4]. We assume that the two impacts are both Ornstein-Uhlenbeck (OU) type stochastic process since we think the instantaneous impact impact has a bit of noise around a deterministic value or average. We explain how the agent would balance the price risk and price impacts when he/she executes a large number of shares by using MOs. The optimal execution strategies with the stochastic price impacts are compared with the results of our benchmark, which is the Almgren-Chriss strategy. We estimate model parameters by using the historical NASDAQ millisecond-stamped messages, and show the performance of strategies.

The breakthrough made by Almgren and Chriss in [2] was regarded as the first one to directly address the issue of permanent and temporary market impact in a continuous time setting, and here Chapter 3 is considered as one of the milestones within the spectrum of high frequency optimal execution study. Since then, the area has attracted consirable attention and numerous extensions have been added to the Almgren-Chriss execution strategy [7]. Cartea's model incorporated the permanent impact of market order-flow in an close-form execution strategy, which consists of an Almgren-Chriss execution strategy adjusted by a weighted-average of the future expected net order-flow [7]. Here, we applied the technique developed by Cartea to solve the new problem of stochastic price impacts.

The rest of this chapter is organised as follows. Section 3.2 gives an overview of order flow and price impacts. Through the empirical data, we show the cross effects between temporary and permanent price impact. In section 3.3, we develop the model for the execution strategy, present the general stochastic process with the order flow and price impacts, and derive the optimal execution strategy. We estimate the parameters in the model and show the performance of the strategy in section 3.4 and conclude in section 3.5.

3.2 Order Flow and Price Impacts

It is generally undesirable to execute large orders in a very limited timespan, because large orders walk the LOB, and as such the average execution price is worse than the best quote posted in the book. A widespread strategy to avoid price impact when executing large orders is to split the order into smaller blocks which are then executed over a time window. This strategy reduces the price impact of the trades to complete the large order, but is exposed to price risk due to fluctuations of the asset's price over the execution window. The risk in price movements may be against the investor's trade direction, upward (resp. downward) pressure in prices if investor is buying (resp. selling), as a result of the one-sided pressure of her MOs over the execution window.

Hypothetically, all market participants' MOs have both temporary and permanent effects on prices. It will be assumed later that both of them are stochastic. In the remaining part of this section, we present statistics and parameter estimated for permanent and temporary price impact for stocks traded in NASDAQ in 2017, as in [10].

Permanent Price Impact

It is supposed that a linear relation between net order-flow which is defined as the difference between the volume of buy and sell MOs and changes in the price. Thus, every trading day we perform the regression

$$\Delta S_n = b\,\mu_n + \varepsilon_n\,, \tag{3.1}$$

where $\Delta S_n = S_{n\tau} - S_{(n-1)\tau}$ is the change in the mid price, μ_n is net order-flow we defined as the difference between the volume of buy and sell MOs during the time

interval $[(n-1)\tau, n\tau]$, ε_n is the error term, and b is the permanent price impact parameter that we wish to estimate. In empirical analysis we choose $\tau = 5\,\text{min}$, during which interval 99% price change were within the range $[-0.1, 0.1]$ with reference to [4].

See [3] for a discussion on linear market impact using proprietary execution data. The first row in Table 3.1 shows the mean and standard deviation of the daily estimate for b by removing the upper and lower 0.5% tails of the data (i.e. winsorize the data) at first and then carrying out a robust linear regression on model (3.1).

Shape of LOB and temporary price impact

	FARO	**INTC**	**NTAP**	**ORCL**	**SMH**
$\widehat{b}$	1.42×10^{-4}	6.17×10^{-7}	5.96×10^{-6}	1.82×10^{-6}	5.48×10^{-6}
	(1.00×10^{-4})	(2.28×10^{-7})	(2.35×10^{-6})	(7.40×10^{-7})	(4.47×10^{-6})
$\widehat{k}$	4.25×10^{-4}	3.61×10^{-7}	3.23×10^{-5}	1.20×10^{-6}	6.49×10^{-6}
	(2.88×10^{-3})	(7.40×10^{-7})	(3.52×10^{-4})	(3.90×10^{-6})	(8.68×10^{-5})
$\frac{\widehat{b}}{\widehat{k}}$	0.82	2.18	1.711.1	2.00	6.90
	(0.65)	(0.63)	(0.72)	(0.70)	(5.60)
Midprice	40.55	23.04	38.33	33.67	37.90
σ	15.1%	3.9%	7.8%	5.4%	6.7%
λ^+	17.00	336.35	305.36	349.44	47.84
	(10.01)	(143.11)	(163.27)	(166.31)	(30.15)
$\mathbb{E}[\eta^+]$	144.40	136.35	309.19	748.57	380.37
	(22.14)	(324.86)	(55.15)	(196.32)	(135.87)
λ^-	17.88	324.99	298.33	336.40	46.76
	(11.58)	(147.24)	(153.46)	(175.46)	(28.92)
$\mathbb{E}[\eta^-]$	104.21	1464.90	313.48	790.67	383.71
	(22.60)	(325.82)	(52.63)	(197.57)	(134.84)

Table 3.1: Permanent and temporary price impact (sell side) parameters for NASDAQ stocks, average volume of MOs, average midprice, σ volatility (annualized) of price returns, hourly mean arrival of MOs $\lambda^{\pm}$, and average volume of MOs $\mathbb{E}[\eta^{\pm}]$. The standard deviation of the estimate is shown in parentheses. Data are from NASDAQ 2017.

Changes in liquidity posted in the LOB are common and unpredictable. There are a number of factors that affect the shape and dynamics of the LOB, including arrival and processing of news, idiosyncratic reasons that prompt market participants to supply liquidity, and how market participants reshuffle their LOs due to the changes

in the LOB and the arrival of MOs. Consequently, liquidity takers cannot predict how their MOs will walk the LOB nor can they quantify (ex-ante) the price impact of their MOs. In this subsection we employ trade data to propose a model of the price impact that MOs have on the book.

We define the price impact generated from walking the LOB as the difference between the cash an investor received from liquidating shares by using an MO and the best bid (the best ask for acquiring shares). We denote the amount of shares sold by the investor at time t by ΔQ_t (i.e. the change in the inventory Q_t), and we assume that the price impact is linear in the size of the order. Specifically, the difference between the best bid and the cash received by the investor is the temporary price impact of $k_t \, \Delta Q_t$, where k_t is the temporary price impact parameter.

To estimate k_t throughout the trading day, we take a snapshot of the buy side of the LOB per second, determine the price per share for various volumes of an MO that walks the LOB, compute the difference between the average price per share and the best bid at that time, and perform a linear regression. The slope of the linear regression is an estimate of the temporary price impact per share. We do this each second of each trading day. Figure 3.1 shows the estimation of the temporary impact parameter for ORCL on 12 April 2017. The left panel shows the entire day and the right panel is a five-minute window of the same day from 11.00 a.m. to 11.05 a.m.

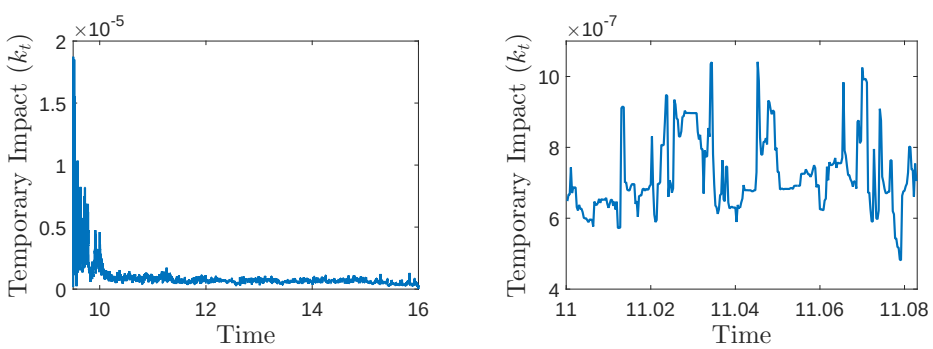

Figure 3.1: An illustration of temporary impacts estimated from the snapshots of the LOBs using ORCL on 12th April, 2017. The right panel shows the estimation of the parameter between 9.30 a.m. and 4.00 p.m. The right panel shows the estimation of the impact parameter impacts from 11.00 a.m. to 11.05 a.m.

The second row in Table 3.1 reports the mean and standard deviation of these daily estimations after we excluded the first and last half-hour of the trading day, and removed the upper and lower 0.5% tails of the data. As discussed in [10], including both sides of the LOB to estimate the impact parameter k does not affect the results; the order of magnitude of the estimated k is statistically the same for both sides of the book.

Table 3.1 also reports the average mid-price, and the (annualized) volatility (σ) of price returns calculated using open-to-close prices and employing five-minute windows (to remove any excess spurious volatility due to micro-structure noise), the same setting as Álvaro Cartea used in his model [4]. Additionally, the average (hourly) number of buy and sell MOs, which is denoted by λ^+ and λ^- respectively, is also reported, as well as the mean volume of MOs, which is denoted by $\mathbb{E}[\eta^+]$ and $\mathbb{E}[\eta^-]$ respectively. For example, in NASDAQ, INTC receives 439 market buy orders on an hourly basis, with an average of $1,049$ shares per order. For both sides of the LOB, the parameter estimations of MO arrival and mean volume are statistically the same. And it is expected to remain for a long time. There could be days or periods of the day where there are more activities on the buy side or sell side, but in the long run buy and sell MOs will be symmetric.

3.2.1 Cross Effects: Temporary and Permanent Price Impact

The analysis above looks at temporary and permanent effects separately, but their joint dynamics are a relevant quantity in execution algorithms. The trade-off between costs that stem from walking the book and permanent impact is taken into account in liquidation and acquisition strategies. Intuitively, increasing (resp. decreasing) execution speeds can expose the strategy to costs from walking the LOB,

but it also can lessen (resp. exacerbates) the effect of future adverse price trends caused by the investor's one-sided pressure on prices.

When both types of impact are linear in rates of trading, this trade-off is partly captured by the ratio b/k. For example, the work of [10] shows how to optimally execute positions when mid-prices are affected by order-flow. The authors show that the optimal speed of execution consists of a deterministic Almgren-Chriss-like strategy plus a term that accounts for expected order-flow, which is proportional to the ratio b/k. A similar result is obtained in [9], which shows how to execute orders that target VWAP (volume-weighted-average-price).

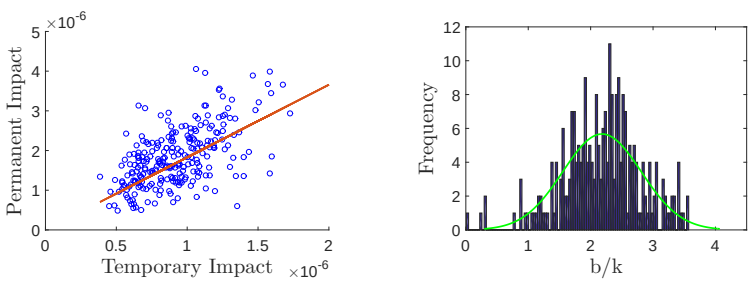

Figure 3.2: Price Impact INTC using daily observations for 2017.

On the left side of Figure 3.2, we show a scatter plot of the daily pair (k, b) for INTC. It is clear that there is a positive relation between temporary and permanent impact. Usually, high (low) permanent impact days are those in which MOs of the same volume must deplete more (less) levels of the book. The right side of the figure depicts a histogram of b/k, which shows that this ratio ranges between 0.5 and 5 and is symmetric around 2.5 approximately.

Finally, Table 3.2 shows the correlation between b and k, and their skewness. Since b/k is rather positive, it is safe to assume b and k have the same sign. Intuitively, one expects these two impacts to be correlated because execution algorithms will trade off permanent and temporary impact.

	FARO	INTC	NTAP	ORCL	SMH
$corr(b,k)$	-0.0532	0.1072	0.0035	-0.0244	-0.0616
$skew(b)$	1.38	1.00	0.69	0.58	1.28
$skew(k)$	15.14	10.30	13.69	15.21	15.73

Table 3.2: Skewness of b and k and their correlation.

3.3 Optimal Execution with Stochastic Price Impact

The investor must choose the speed at which he/she sends MOs to liquidate $\mathfrak{N} > 0$ shares over a trading horizon $T > 0$. Here we focus on the liquidation problem – the setup for the acquisition problem is similar. We denote the liquidation speed, which is controlled by the investor, by $\nu = (\nu_t)_{\{0\leq t\leq T\}}$, and denote the controlled inventory by $Q^\nu = (Q^\nu_t)_{\{0\leq t\leq T\}}$, which is affected by how fast he/she trades, and satisfies

$$dQ^\nu_t = -\nu_t\, dt\,, \qquad\qquad Q^\nu_0 = \mathfrak{N}\,. \tag{3.2}$$

Mid-price dynamics and price impact. The mid-price process $S^\nu = (S^\nu_t)_{\{0\leq t\leq T\}}$ satisfies the SDE

$$dS^\nu_t = b_t\,(\mu_t - \nu_t)\, dt + \sigma\, dW^1_t\,, \qquad S^\nu_0 = S\,, \tag{3.3}$$

where $\mu_t = \mu^+_t - \mu^-_t$ is the net order flow of market participants. The participants buy MOs at speed $\mu^+ = \{\mu^+_t\}_{0\leq t\leq T}$ and sell MOs at speed $\mu^- = \{\mu^-_t\}_{0\leq t\leq T}$, which exclude the investor's own trading rate. They are jointly assumed Markov, cadlag and bounded $\mathbb{P}-$a.s.

The shocks to the mid-price are represented by the standard Brownian motion $W^1 = (W^1_t)_{\{0\leq t\leq T\}}$, and we assume that the buy and sell order flows $\mu^\pm$ are independent of W^1. The process $b = (b_t)_{\{0\leq t\leq T\}}$ represents the permanent price impact that order-flow has on mid-prices, which we will provide further details below.

The execution price received by the investor is

$$\widehat{S}_t^\nu = S_t^\nu - \left(\tfrac{1}{2}\Delta + k_t\,\nu_t\right)\,, \tag{3.4}$$

where $\Delta \geq 0$ is the bid-ask spread, which is assumed to be invariable. Here we only consider the problem of liquidating the assets, thus we set $\Delta = 0$ and let S_t^ν be the bid-price. We assume the bid-price have the same dynamic as the mid-price. The temporary price impact process $k = (k_t)_{\{0\leq t\leq T\}}$ satisfies the stochastic differential equation (SDE)

$$dk_t = \beta\,(\xi - k_t)\;dt + \sigma_k\,dW_t^2 + d\sum_{i=1}^{N_t}\eta_i\,. \tag{3.5}$$

The parameters $\beta > 0$, $\sigma_k > 0$, and $\xi \geq 0$, are all constants, $W^2 = (W_t^2)_{\{0\leq t\leq T\}}$ is a standard Brownian motion, $N = (N_t)_{\{0\leq t\leq T\}}$ is a Poisson process with intensity λ_k, and $\{\eta\} \overset{i.i.d.}{\sim} F$, where F is a distribution function with support on $[0,\infty)$ and $\mathbb{E}[\eta] = \eta_0 < \infty$, where $\mathbb{E}[\,\cdot\,]$ denotes the expectation operator.

We assume the permanent price impact parameter $b = (b_t)_{\{0\leq t\leq T\}}$ satisfies

$$b_t = \ell_1 + \ell_2\,k_t\,, \tag{3.6}$$

where ℓ_1 and ℓ_2 are constants, see Figure 3.2 and analysis therein.

The processes N, W^1, W^2, $\mu^\pm$, and random variable η are all independent of each other.

Performance criterion and value function. The investor's objective is to maximize the cash proceeds from selling the shares. We denote the controlled cash process by $X^\nu = (X_t^\nu)_{\{0\leq t\leq T\}}$, which satisfies the SDE

$$dX_t^\nu = \widehat{S}_t^\nu\,\nu_t\,dt\,, \qquad X_0^\nu = x\,. \tag{3.7}$$

The investor's performance criterion is

$$H^{\nu}(t,x,S,b,k,\boldsymbol{\mu},q)=\mathbb{E}_{t,x,S,b,k,\boldsymbol{\mu},q}\left[X_T+Q_T^{\nu}\left(S_T^{\nu}-\tfrac{1}{2}\Delta-\alpha\,Q_T^{\nu}\right)-\phi\int_t^T(Q_u^{\nu})^2\,du\right], \tag{3.8}$$

where $\boldsymbol{\mu}=\{\mu^+,\mu^-\}$, and the operator $\mathbb{E}_{t,x,S,k,b,\boldsymbol{\mu},q}[\,\cdot\,]$ represents expectation in the condition (with a slight abuse of notation) of $X_t=x$, $S_{t^-}=S$, $b_{t^-}=b$, $k_{t^-}=k$, $\mu^+_{t^-}=\mu^+$, $\mu^-_{t^-}=\mu^-$ and $Q_t=q$, and its value function is

$$H(t,x,S,b,k,\boldsymbol{\mu},q)=\sup_{\nu\in\mathcal{A}}H^{\nu}(t,x,S,b,k,\boldsymbol{\mu},q)\,, \tag{3.9}$$

where $\mathcal{A}$ is the set of admissible strategies consisting of $\mathcal{F}$-predictable processes such that $\int_0^T|\nu_u|\,du<+\infty$, $\mathbb{P}-$ a.s..

The right side of the performance criterion (3.8) contains three terms. The first term X_T is the investor's terminal cash from liquidating the shares throughout the trading horizon. The second term is the proceeds received by the investor from liquidating any remaining inventory Q_T^{ν} at the end of the strategy. This leftover inventory is liquidated at bid-price S_T^{ν}, then pays the costs associated with crossing the spread, liquidity taking fees, and market impact. All these costs are captured by the liquidation penalty parameter $\alpha\geq 0$.

Finally, the third term is the running penalty $\phi\int_t^T(Q_u^{\nu})^2\,du$ where $\phi\geq 0$ is the inventory penalty parameter. This penalty does not affect the investor's revenues, but affects the optimal liquidation rate. When the value of the inventory penalty parameter ϕ is high then carrying inventory becomes very expensive from the utility point of view. As such, the liquidation speed will be higher. Therefore, this parameter could be interpreted as an urgency parameter. Involving the running inventory penalty is also justified in a setting where the agent considers model uncertainty, i.e. he/she is ambiguity aversion. Álvaro Cartea shows that including the running penalty is equivalent to the agent considering alternative models with stochastic

drifts [6]. But it penalizes those models using relative entropy. Under that circumstance, the higher the value of ϕ is, the less confident about the trend of the bid-price the agent will be.

3.3.1 Dynamic Programming Equation

The dynamic programming equation associated with the optimal control problem (3.9) suggests that the value function $H(t, x, S, b, k, \boldsymbol{\mu}, q)$ is the unique solution of the Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} \partial_t H + \frac{1}{2}\sigma^2\,\partial_{SS}H - \phi\,q^2 + \mathcal{L}^{\boldsymbol{\mu}}H + \mathcal{L}^{k}H + b\,\boldsymbol{\mu}\,\partial_S H \\ + \sup_{\nu}\left[(S - k\,\nu)\,\nu\partial_x H - b\,\nu\,\partial_S H - \nu\,\partial_q H\right] &= 0\,, \end{aligned} \tag{3.10}$$

where

$$\mathcal{L}^{k}H = \beta\,(\xi - k)\,\partial_k H + \frac{1}{2}\,\sigma_k^2\,\partial_{kk}H + \lambda\,\mathbb{E}\left\{[H(t, x, S, k + \eta, q) - H(t, x, S, k, q)]\right\}\,,$$

and $\mathcal{L}^{\boldsymbol{\mu}}$ is the generator of the process $\boldsymbol{\mu}$.

Upon taking the sumpremum in the HJB (3.10), we obtain the optimal speed of trading in feedback form as:

$$\nu^* = \frac{S}{2k} - \frac{b\,\partial_S H + \partial_q H}{2\,k\,\partial_x H}\,,$$

and the HJB becomes

$$\partial_t H + \frac{1}{2}\sigma^2\,\partial_{SS}H - \phi\,q^2 + \mathcal{L}^{\boldsymbol{\mu}}H + \mathcal{L}^{k}H + b\,\boldsymbol{\mu}\,\partial_S H + \frac{1}{4k\,\partial_x H}\left[S\,\partial_x H - (b\,\partial_S H + \partial_q H)\right]^2 = 0\,. \tag{3.11}$$

Since in closed-form the HJB (3.10) is not able to be solved, we will employ numerical methods to obtain the investor's optimal liquidation speed. Correspondingly

we assume the ansatz:

$$H(t,x,S,b,k,\boldsymbol{\mu},q) = x + q\,S + h_0(t,b,k,\boldsymbol{\mu}) + h_1(t,b,k,\boldsymbol{\mu})\,q + h_2(t,b,k,\boldsymbol{\mu})\,q^2 \ ,$$

subject to terminal conditions $h_0(T,b,k,\boldsymbol{\mu}) = h_1(T,b,k,\boldsymbol{\mu}) = 0$ and $h_2(T,b,k,\boldsymbol{\mu}) = -\alpha$. And the optimal liquidation speed in the feedback form is

$$\boxed{\nu^* = -\frac{b\,q + h_1 + 2\,h_2\,q}{2\,k}}\,. \tag{3.12}$$

Upon substituting it back into the HJB (3.10) we obtain:

$$\begin{aligned}\Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_0 + \Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_1\,q + \Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_2\,q^2 & \\ -\phi\,q^2 + b\,\boldsymbol{\mu}\,q + \frac{1}{4k}(b\,q + h_1 + 2\,h_2\,q)^2 &= 0\end{aligned} \tag{3.13}$$

Recalling that $b_t = \ell_1 + \ell_2\,k_t$ and collecting like terms in q leads to the following coupled system of PIDEs:

$$\Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_0 + \frac{1}{4k}\,h_1^2 = 0\ , \tag{3.14}$$

$$\Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_1 + (\ell_1 + \ell_2\,k)\,\boldsymbol{\mu} + \frac{1}{2k}\,h_1\,(\ell_1 + \ell_2\,k + 2\,h_2) = 0\ , \tag{3.15}$$

$$\Big(\partial_t + \mathcal{L}^{\boldsymbol{\mu}} + \mathcal{L}^k\Big)\,h_2 - \phi + \frac{1}{4k}\,(\ell_1 + \ell_2\,k + 2h_2)^2 = 0\ . \tag{3.16}$$

We show the numerical scheme that we employed to solve this coupled system of PIDEs. See Appendix 7.1. However, note that if

$$\beta = \xi = 0, \quad \sigma_k = 0\ , \quad \lambda_k = 0\,,$$

in the temporary price impact model (3.5), and in the meanwhile we assume the

temporary and permanent impact parameters are constants to find a solution, i.e.

$$k_t = k^C\,, \qquad \text{and} \qquad b_t = \ell_1 + \ell_2\, k^C = b^C\,. \tag{3.17}$$

In this case, the coupled system of PIDEs becomes

$$\begin{aligned}
(\partial_t + \mathcal{L}^{\boldsymbol{\mu}})\, h_0 + \frac{1}{4\,k^C}\, h_1^2 &= 0\,, &(3.18)\\
(\partial_t + \mathcal{L}^{\boldsymbol{\mu}})\, h_1 + b^C\, \boldsymbol{\mu} + \frac{1}{2\,k^C}\, h_1\,(b^C + 2\,h_2) &= 0\,, &(3.19)\\
(\partial_t + \mathcal{L}^{\boldsymbol{\mu}})\, h_2 - \phi + \frac{1}{4\,k^C}\, \Big(b^C + 2\,h_2\Big)^2 &= 0\,, &(3.20)
\end{aligned}$$

with the terminal conditions $h_0(T, \boldsymbol{\mu}) = h_1(T, \boldsymbol{\mu}) = 0$ and $h_2(T, \boldsymbol{\mu}) = -\alpha$, which can be solved in closed-form.

To solve h_2, we note that equation (3.20) is of Riccati type without source terms in $\boldsymbol{\mu}$, and its terminal condition is independent of $\boldsymbol{\mu}$, hence the solution must be independent of $\boldsymbol{\mu}$. Accordingly, equation (3.20) can be integrated accurately. First, let $h_2(t) = \chi(t) - \frac{1}{2}\, b^C$, then rearranging equation (3.20). We obtain

$$\frac{\partial_t \chi}{k\,\phi - \chi^2} = \frac{1}{k}\,, \tag{3.21}$$

subject to $\chi(T) = -\alpha + \frac{1}{k^C}$.

Next, integrate both sides of the above from t to T:

$$\chi(t) = \sqrt{k^C\,\phi}\frac{1 + \zeta\, e^{2\,\gamma(T-t)}}{1 - \zeta\, e^{2\,\gamma(T-t)}}\,, \tag{3.22}$$

where $\gamma = \sqrt{\frac{\phi}{k^C}}$ and $\zeta = \frac{\alpha - \frac{1}{2}\,k^C + \sqrt{k^C\,\phi}}{\alpha - \frac{1}{2}\,k^C - \sqrt{k^C\,\phi}}$.

Recalling that $h_2(t) = \chi(t) - \frac{1}{2}\, b^C$ and substituting the equation (3.22) back into it, we find

$$h_2(t) = \sqrt{k^C\,\phi}\,\frac{1 + \zeta\, e^{2\,\gamma(T-t)}}{1 - \zeta\, e^{2\,\gamma(T-t)}} - \frac{1}{2}\, b^C\,, \tag{3.23}$$

where γ and ζ are defined as above.

We then move on to solve (3.19), which is a linear PIDE satisfied by h_1, where $b^C\,\boldsymbol{\mu}$ is a source term and $h_2 + \frac{1}{2}\,b^C$ acts as an effective discount rate. The solution of this PIDE equation can be derived by Feynman-Kac, thus

$$h_t(t,\boldsymbol{\mu}) = b\,\mathbb{E}_{t,\boldsymbol{\mu}}\left[\int_t^T \exp\left\{\frac{1}{k}\int_t^u \left(h_2(s) + \frac{1}{2}\,b\right)\,ds\right\}\boldsymbol{\mu}_u du\right]\,,$$

which can be simplified to

$$h_1(t,\boldsymbol{\mu}) = b\int_t^T \left(\frac{e^{-\gamma(T-u)} - \zeta\,e^{\gamma(T-u)}}{e^{-\gamma(T-t)} - \zeta\,e^{\gamma(T-t)}}\right)\mathbb{E}_{t,\boldsymbol{\mu}}[\boldsymbol{\mu}_u]du\,. \tag{3.24}$$

Similarly, by the Feynman-Kac theorem

$$h_0(t,\boldsymbol{\mu}) = \frac{1}{4\,k}\int_t^T \mathbb{E}_{t,\boldsymbol{\mu}}[h_1^2(t,\boldsymbol{\mu}_u)]\,du\,. \tag{3.25}$$

Putting the above results together, we find that the optimal trading speed is given by

$$\nu_t^* = \gamma\,\frac{\zeta\,e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta\,e^{\gamma(T-t)} - e^{-\gamma(T-t)}}\,Q_t^{\nu^*} - \frac{b^C}{2\,k^C}\int_t^T \left(\frac{\zeta\,e^{\gamma(T-u)} - e^{-\gamma(T-u)}}{\zeta\,e^{\gamma(T-t)} - e^{-\gamma(T-t)}}\right)\mathbb{E}[\boldsymbol{\mu}_u\,|\,\mathcal{F}_t^{\boldsymbol{\mu}}]\,du\,, \tag{3.26}$$

which is also the optimal liquidation speed derived in [10]. Here $\mathcal{F}_t^{\boldsymbol{\mu}}$ denotes the natural filtration generated by $\boldsymbol{\mu}$. If we have $\zeta \to 1$, and the optimal trading speed simplifies to

$$\lim_{\zeta\to 1}\nu_t^* = \gamma\,\frac{\cosh(\gamma(T-t))}{\sinh(\gamma(T-t))}\,Q_t^{\nu^*} - \frac{b^C}{2\,k^C}\int_t^T \left(\frac{\sinh(\gamma(T-u))}{\sinh\gamma(T-t))}\right)\mathbb{E}[\boldsymbol{\mu}_u\,|\,\mathcal{F}_t^{\boldsymbol{\mu}}]\,du\,, \tag{3.27}$$

3.4 Performance of Strategy

In this section we employ simulations to illustrate the performance of the execution strategy. To this end, in subsection 3.4.1 we estimate the model parameters for price impact, which is followed by subsection 3.4.2. In that part, we will discuss the performance of the strategy and compare it to the Almgren-Chriss (AC) optimal execution strategy.

3.4.1 Estimation of Model Parameters

We employ a Maximum Likelihood Estimation (MLE) to estimate the parameters of the temporary price impact which appears in the SDE (3.5). We minimize the negative of the log-likelihood

$$\log L = -\sum \left\{ \log \left[\lambda \Delta t \, \phi \left(\zeta_{1,t}, \sqrt{\eta^2 + \sigma^2)} \right) + (1 - \lambda \Delta t) \, \phi \left(\zeta_{2,t}, \sigma \right) \right] \right\} \ ,$$

where the processes $\gamma_{1,t}$ and $\gamma_{2,t}$ are defined as

$$\begin{aligned} \zeta_{1,t} &= \Delta k_t - \beta \left(\xi - k_{t-1} \right) - \eta_t \, , \qquad \text{(when a jump occurs)} \, , \\ \zeta_{2,t} &= \Delta k_t - \beta \left(\xi - k_{t-1} \right) , \qquad \text{(when no jump occurs)} \, . \end{aligned}$$

To obtain the initial values that we employ in the MLE above, we assume that there are no jumps in the temporary impact parameter. Consequently, (3.5) becomes a standard Orstein-Uhlenbeck process. And using standard MLE methods, we will obtain an (initial) estimate of the parameters β, ξ, and σ_k. Next, to obtain the initial value for the mean jump size used in the MLE we assume any increase in the impact parameter k_t is caused by a jump, and employ these observations to calculate an initial average jump size.

Once we obtained the parameters of the SDE satisfied by k, we would use ordinary

least squares to estimate the parameters of the model for b.

Table 3.3 shows the results of the parameter estimates. First, the daily average temporary and permanent impact we got indicate a hidden linear relation, which is the estimation of $\widehat{\ell}_1$ and $\widehat{\ell}_2$. Then, in each trading day, we have an estimation set of $\widehat{\beta}$, $\widehat{\xi}$, $\widehat{\sigma}_k$, $\widehat{\eta}_0$ and $\widehat{\lambda}_k$. The number showed in the first part of the table (first five rows) is the mean and standard deviation of each trading day during one year respectively.

	FARO	INTC	NTAP	ORCL	SMH
$\widehat{\beta}$	459.91	759.29	1610.71	2418.90	5004.40
	(252.19)	(468.45)	(1054.57)	(1958.69)	(6215.44)
$\widehat{\xi}$	1.86×10^{-4}	5.66×10^{-7}	2.81×10^{-5}	1.83×10^{-6}	2.74×10^{-6}
	(1.17×10^{-4})	(2.65×10^{-6})	(3.82×10^{-4})	(1.19×10^{-5})	(2.50×10^{-5})
$\widehat{\sigma}_k$	4.50×10^{-3}	1.29×10^{-5}	2.65×10^{-4}	6.63×10^{-5}	1.62×10^{-4}
	(2.70×10^{-3})	(7.12×10^{-5})	(1.92×10^{-4})	(4.82×10^{-45})	(4.38×10^{-4})
$\widehat{\eta}_0$	6.22×10^{-5}	1.85×10^{-8}	5.16×10^{-7}	8.39×10^{-8}	3.00×10^{-7}
	(6.63×10^{-5})	(3.73×10^{-8})	(7.23×10^{-7})	(4.09×10^{-8})	(1.04×10^{-6})
$\widehat{\lambda}_k$	1100	7238	6336	6937	2617
	(358.84)	(881.40)	(944.36)	(987.79)	(824.79)
$\widehat{\ell}_1$	1.42×10^{-4}	6.05×10^{-7}	5.96×10^{-6}	1.82×10^{-6}	5.51×10^{-6}
	(6.38×10^{-6})	(1.59×10^{-8})	(1.49×10^{-7})	(4.89×10^{-8})	(2.83×10^{-7})
$\widehat{\ell}_2$	-1.85×10^{-3}	0.0330	2.33×10^{-5}	-4.63×10^{-3}	-3.18×10^{-3}
	(2.20×10^{-3})	(0.0194)	(4.23×10^{-4})	(0.012)	(3.26×10^{-3})

Table 3.3: Parameter estimates for temporary and price impact models. Data are from all trading days in NASDAQ 2017.

3.4.2 Simulations of the Strategy with Both Impacts

In this section, we perform simulations to show the behaviour of the optimal strategy by using our model. We use the data of INTC as an example. The performance of other stocks can be found in the Appendix 7.2. Using the linear regression results shown in the table 3.3, it can be seen that the relation between temporary and permanent impacts is

$$b_t = 6.05 \times 10^{-7} + 0.033\, k_t \ , \tag{3.28}$$

i.e. $\widehat{\ell}_1 = 6.05 \times 10^{-7}$ and $\widehat{\ell}_2 = 0.033$. Meanwhile, we set the price of the trading asset at the starting point as $S_0 = 23.04$ and its volatility as $\sigma = 0.039$. Set the other parameters as follows.

$$\beta = 759.29 \ , \quad \xi = 5.66 \times 10^{-7} \ , \quad \sigma_k = 1.29 \times 10^{-5} \ , \quad \lambda_k = 7238 \ ,$$
$$\mathbb{E}[\eta] = \eta_0 = 1.85 \times 10^{-8} \ , \quad \alpha = 10^3 \, \xi \ , \quad \phi = 10^3 \, \xi \ .$$

Presumably the investor must liquidate $\mathfrak{N} = 1000$ shares of the stock INTC over $T = 1$ day. And we set the net order flow as zero for short, i.e $\boldsymbol{\mu} = 0$.

If the net order flow is zero, the optimal strategy in (3.26) will change to the Almgren-Chriss strategy in Chapter 2.

We use this trading rate as our comparison benchmark, and compare the financial performance of the strategy developed here with that of an agent who employs the Almgren-Chriss liquidation strategy without incorporating stochastic temporary and permanent price impacts. The temporary price impact in the benchmark is $k^C = \xi$ (recall that ξ is the long term level of k_t, see (3.5)). Meanwhile, the permanent impact in the benchmark is $b^C = \ell_1 + \ell_2 \ k^C$. And we perform 10,000 simulations.

3.4.3 Performance comparison of Both Stochastic Impacts

The top panel of Figure 3.3 shows two paths of temporary and permanent impact. The bottom panel shows one simulation path for inventory and liquidation speed. Here, we show such a changing path compared with the Almgren-Chriss strategy. We can see that when the temporary impact is big, the trading speed will be slow and the inventory will reduce more smoothly. The trading speed decreases as the inventory gets smaller and smaller.

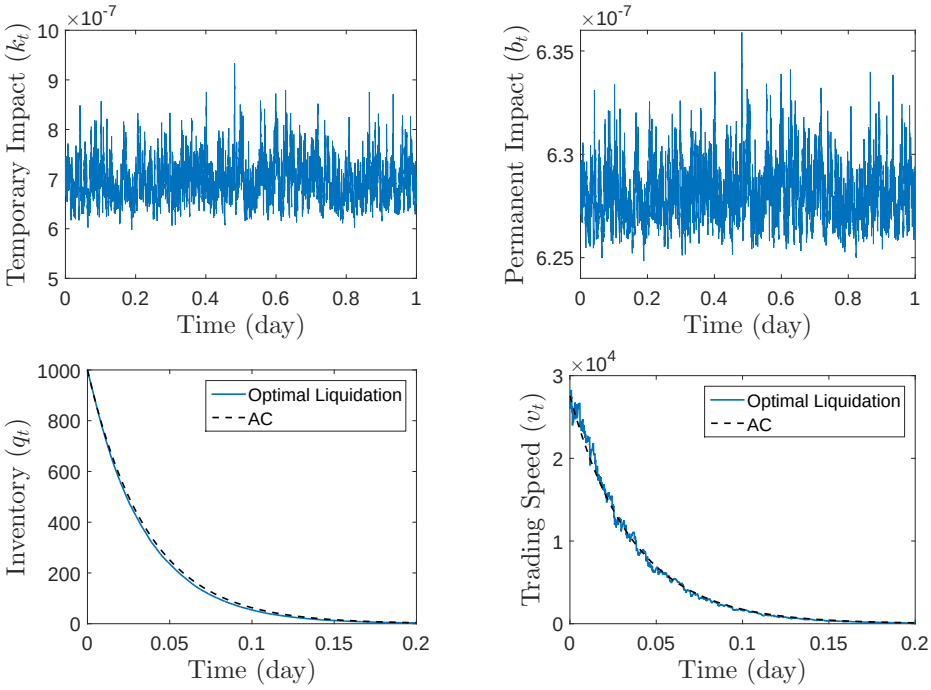

Figure 3.3: Optimal trading with stochastic temporary and permanent impact

The top panel of Figure 3.3 shows two paths of temporary and permanent impact. The bottom panel shows one simulation path for inventory and liquidation speed during the first 78 minutes after market opening (0.2 day = 0.2 * (16 - 9.5) * 60 = 78 minutes). Here, we show such a changing path compared with the Almgren-Chriss strategy. We can see that when the temporary impact is big, the trading speed will be slow and the inventory will reduce more smoothly in the bottom right figure. The trading speed decreases as the inventory gets smaller and smaller.

To compare the performance of the strategy with the performance obtained by employing Almgren-Chriss, we measured this value by basis points using

$$\frac{X_T^{\nu^*} - X_T^{AC}}{X_T^{AC}} \times 10^4\,, \tag{3.29}$$

where $X_T^{v^*}$ and X_T^{AC} mean our terminal cash and the one using Almgren-Chriss, which are obtained by the agent in the end. These values include the penalty.

We calculate the mean and the percentage of the runs where the strategy of

stochastic temporary impact underperforms those of constant impact. We list the results in the Figure 3.4 and Table 3.4.

With stochastic impacts, the performance is expected to be better than that of a constant case. The mean of the performance is 13.8281. And no runs show that the strategy with stochastic impacts underperforms the strategy in a constant case. The mean of the performance shows us that in both scenarios, the terminal cash in the stochastic case tends to be larger than that in the constant ones. According to Table 3.4, the non-constant strategy outperforms the constant one in 100% of the simulations.

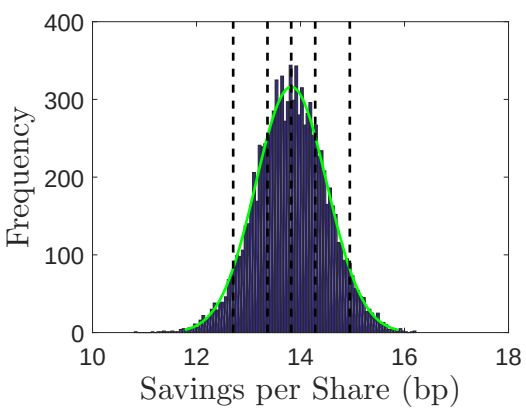

Figure 3.4: The savings per share measured in basis points

mean	13.8210
stdev	0.0897
5%	12.7048
25%	13.3643
50%	13.8278
75%	14.2824
95%	14.9475
$X_T^{v^*} < X_T^C$	0%

Table 3.4: Quantiles of relative performance in basis points

3.4.4 Different levels of running inventory penalty

In the previous comparison, we set up the inventory penalty as $\phi = 10^3\,\xi$. Next, we will list some different levels of running inventory penalty to check whether it

can influence the performance. Table 3.5 shows that as the penalty increases, the mean of performance will grow too, at a rather impressive speed. The penalty helps to improve the performance and decrease the standard deviation, which means the Sharpe ratio increases.

ϕ	FARO	INTC	NTAP	ORCL	SMH
10ξ	0.9688	1.2207	1.3893	1.4424	1.4479
	(0.4384)	(0.2232)	(0.0632)	(0.0079)	(0.0022)
$10^2\xi$	3.7165	4.3678	4.8294	4.9712	4.9866
	(0.4030)	(0.1459)	(0.0406)	(0.0060)	(0.0015)
$10^3\xi$	11.7752	13.8210	15.2839	15.7334	15.7817
	(0.3162)	(0.0897)	(0.0259)	(0.0049)	(0.0023)
$10^4\xi$	37.3033	43.8171	48.4373	49.8660	50.0197
	(0.2317)	(0.0456)	(0.0181)	(0.0056)	(0.0043)
$10^5\xi$	118.5907	139.4379	154.2430	158.8259	159.3188
	(0.1619)	(0.0258)	(0.0156)	(0.0081)	(0.0079)

Table 3.5: Performance comparison with different levels of Inventory Penalty

3.5 Chapter Conclusions

We propose a model that captures the dynamics of temporary and permanent price impact. In this case, temporary price impact refers to a process where MOs walk through the different levels of LOBs which will subsequently be replenished after certain timespan (milliseconds). With regard to permanent price impact, it refers to changes in the permanent bid-price due to the information conveyed by the MOs which are impounded in the price of the asset.

Data from the NASDAQ exchange are employed to motivate a Ornstein-Uhlenback type model which has a jump diffusion. In this model, jumps are achieved according to a Poisson counting process. Meanwhile, the impacts decay to a 'long term' level. We use this to pose an optimal execution problem and employ numerical methods to solve the HJB equation related.

We illustrate the performance of the strategy through taking in simulations and comparing its results with those of the classical Almgren-Chriss model. Particularly, we estimate the model parameters by analysing NASDAQ millisecond-stamped messages for FARO, INTC, NTAP, ORCL, SMH and examine the performance of the strategies through simulations under different scenarios. And it shows that accounting for the stochastic nature of price impact can improve the performance of trading algorithms.

Chapter 4

Stochastic Latency Impact

4.1 Problem Introduction

In this chapter we define latency as the delay time between the submission of the signal and the receipt of the electronic exchange response to that signal, or in another way the time it takes for the signal to travel in the automated trading system. A trader who seeks to optimally liquidate or acquire shares in the electronic exchange will inevitably encounter the latency regardless of the nature of the trading. Latency is vital to traders, as it has an impact on the execution price. This price impact is stochastic because the price of the asset fluctuates during the period of the latency, making the price that was initially observed by the trader different from the one that was ultimately received and executed by the trading system, even though the period of latency can be as short as thousandths and millionths of seconds.

According to our definition, latency is mainly attributed to the limitation of the exchange technology. Here, we do not discuss how to reduce the latency impact, but rather we look for an optimal execution strategy for traders, with stochastic latency impact taken into account. To this end, we assume that all traders are in the same technical environment, and latency is a stochastic factor that always exists in the system. We also assume that traders only use Market Orders (MOs) to do the

execution so as to avoid extra price risk.

Latency may give rise to stochastic fluctuations, making the execution price higher or lower than the observed price that a trader intends to capture when he/she submits the order.

We model latency as Ornstein-Uhlenbeck (OU) type process with jumps, and we employ another OU process to formulate the latency impact as it follows an auto-regression model in the discrete form. We will provide an explicit closed-form expression for the optimal execution strategy, taking stochastic latency into account. The optimal trading speed in our strategy is generated by a dynamic programming problem and is found to be affected by the latency impact in a linear form. Here, we again employ the technique mentioned in Chapter 2 and use it to solve the problem.

The rest of the chapter is organised as follows: Section 4.2 defines the latency in detail. In section 4.3, we set up a model for the execution strategy with the latency impact formulated as an OU process, then we derive the optimal execution strategy, accompanied with stochastic latency, from the model. Finally, in section 4.4, we employ simulations to showcase the performance of the strategy and section 4.5 concludes the chapter.

4.2 Latency and Latency Impact

4.2.1 Latency Impact

We symbolise the price that a trader observes in the the market as S_t. However, the price received eventually by the exchange is different from S_t due to latency. We label the price received by the exchange market as $\widetilde{S}_t$. In this section, we will present the expected difference between the submission of the observed price S_t and

the execution price $\widetilde{S}_{t^+}$, which we define as the latency impact η_t:

$$\eta_t = \mathbb{E}_t \left[S_t - \widetilde{S}_{t^+} | \mathcal{F}_t \right] \ , \tag{4.1}$$

where η_t is an $\mathcal{F}_t$-adapted process based on a filtered probability space$(\Omega \, , \mathcal{F} \, , \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Meanwhile, η_t can be positive or negative, depending on the direction of the price fluctuation.

4.2.2 Stochastic Latency

Based on the definition of latency, namely the delay time between the signal and the response time it takes for the signal to travel in the automated trading system, we make the assumption that latency is given by the stochastic process $L = (L_t)_{\{0\leq t\leq T\}}$. It is a stochastic process that satisfies the stochastic differential equation (SDE)

$$dL_t = \psi \left(\theta - \ L_t \right) \, dt + d \sum_{i=0}^{N_t} J_i \, , \tag{4.2}$$

where θ is the mean value of delay time, ψ is the speed at which latency reverts to its mean level, J are i., i.e. random variables which denote the jump size and follow a uniform distribution $\mathcal{U}(0, J_{max})$, $N = (N_t)_{\{0\leq t\leq T\}}$ is a homogeneous Poisson counting process with arrival intensity λ, independent of J.

Based on this delay time, we simulate latency impact using real data and the Alphabet Inc. stock (GOOG) as an example. When we use a high frequency trading strategy which means trading is executed in every 20ms, and we assume that the best condition is 10ms delay for getting the signal. Here, the best condition means that we need 10ms to send the signal at least due to the objective factors. Therefore, we set $\psi = 0.9$, $\theta = 10$, $\lambda = 1$ and $J_{max} = 20$.

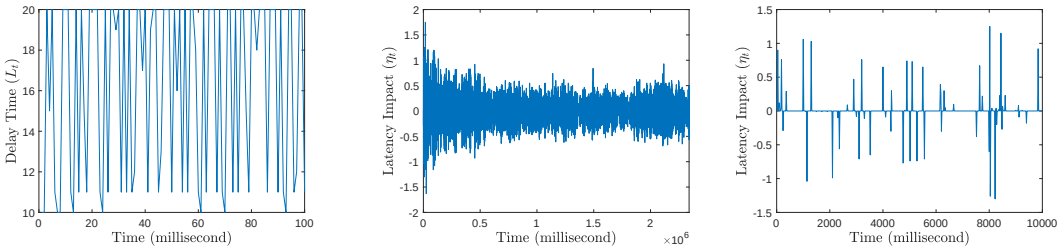

Figure 4.1: The Figure on the left shows the delay time. The center and right figures illustrate the latency impact η_t during one day and a 10000ms time window.

The first graph in the Figure 4.1 is a simulation of the delay time in a 100ms period. We find that jumps close to 20ms present a longer delay time. The shortest duration is 10ms delay time. The graph in the middle shows the latency impact during a single trading day (12th April, 2017). The graph on the right zooms in the latency impact during a 10000ms period. We observed that the behaviour of the latency impact looked like an OU process. This fact encourages us to use an OU type model to describe the latency impact.

4.2.3 Latency Impact Auto-regression Test

The process of latency impact appears to behave like an OU process. We then run the Augmented Dickey-Fuller(ADF) test to check whether the latency series fits the AR(1) model or not.

We write the latency process in discrete form

$$\eta_t = \omega\, \eta_{t-1} + \epsilon_t \ , \tag{4.3}$$

where ω is auto-regression lag one term coefficient and ϵ_t is a white noise process.

We make the null hypothesis that $\omega = 0$ against the alternative hypothesis that $\omega \neq 0$. The test statistic is

$$DF_t = \frac{\widehat{\omega} - 1}{SE(\widehat{\omega})} \ . \tag{4.4}$$

Ticker	AAPL	AMZN	FB	GOOG	INTC	MSFT	NFLX	NVDA	TSLA
mean(η_t)	2.6e-5	9.2e-5	7.3e-6	-1.5e-4	-5.6e-6	1.6e-5	2.2e-6	2.8e-5	-2.5e-4
Std(η_t)	0.0058	0.0673	0.0058	0.0685	0.0035	0.0042	0.0089	0.0080	0.0416
$\widehat{\omega}$	0.1228	0.0830	0.1314	0.0660	0.1285	0.1459	0.0900	0.1033	0.831
	(3.4e-5)	(0.0045)	(3.3e-5)	(0.0447)	(1.2e-5)	(1.7e-5)	(7.9e-5)	(6.3e-5)	(0.0018)
DF_t	1	1	1	1	1	1	1	1	1

Table 4.1: Latency Data Description: This is the case of $\psi = 0.9$, $\theta = 10$, $\lambda = 1$ and $J_{max} = 20$ in the delay time model. All of the parameters are estimated by using the nine stocks traded in NASDAQ. All of the data are taken from 12th April, 2017, in the high frequency form.

If $DF_t = 1$, we reject the null hypothesis, i.e. we conclude that the series is an auto-regressive process.

In Table 4.1, we trade every 10ms as above. We observe the mean of latency impact is very close to zero although the standard deviations are more significant. Here we use the Yule-Walker Method to do the auto regression for the latency impact. The estimated parameters for ω can be found in the table. The numbers in brackets are the mean square errors (MSE) of $\widehat{\omega}$.

The results of the ADF test shows that we should reject the null hypothesis for every ticker. Hence, the latency impact is not a white noise series; there exists auto correlation in the series itself. This leads us to use an OU process to model the latency impact.

We also explore the Brownian Motion case of the stock price, i.e. we test whether the latency impact still fits the AR(1) model in the event that the stock follows a Brownian motion and whether the delay time is based on model 4.2. And we find it still fits the Brownian price. More details can be found in the Appendix.

4.2.4 Robustness Test of Delay Time Model

Here, we check how different values of θ impose influence to the latency impact. The parameter θ depends on the travel distance of the signal. By the lowest latency, trading engines are located physically closer to the exchanges, or even in the same

building (co-location) to further reduce latency. These investors may gain the 'ultra low latency', which means the delay time is below 1ms. In this case, we set $\theta = 0$ to show the best location condition. Normally speaking, the trading engine down the road from the exchange with a distance of 100 miles will have around 1ms of the delay time. For example, the investors in Chicago (around 800 miles to NY) and Los Angeles (around 2800 miles to NY) who both want to trade the stocks on the NYSE will suffer different delay times.

In this section, we keep the other parameter constantly, i.e. $\psi = 0.9$, $\lambda = 1$ and $J_{max} = 20$. We choose $\theta = 0$, $\theta = 5$, $\theta = 10$ and $\theta = 15$ to see whether the latency impact still fits the AR(1) model.

Ticker: GOOG	$\theta = 0$	$\theta = 5$	$\theta = 10$	$\theta = 15$
mean(η_t)	-1.3432e-04	-4.5941e-5	-1.4868e-04	-8.3804e-5
Std(η_t)	0.0504	0.0617	0.0685	0.0734
$\widehat{\omega}$	0.0571	0.0616	0.0660	0.0882
	(0.0025)	(0.0038)	(0.0447)	(0.0053)
DF_t	1	1	1	1

Table 4.2: Different Levels of θ

From table 4.2, we confirm that the latency impact fits AR(1) model. When θ increases, the mean of the latency impact will not change much although volatility of latency will increase. Longer average delay time leads to more uncertainties of the execution price.

4.3 The Model

4.3.1 Stochastic Latency Impact Model

We have the basic model setup for the bid-price (for the liquidation problem, and bid-price for the acquisition problem)of an underlying asset. We adapt this model by including a stochastic impact from the latency. This latency impact follows the OU process with mean zero, which means we do not take the location factor into

consideration and assume that the investors' trading engines here are placed in the same premises as an exchange server. We call this co-location and denote this best condition as the ultra low latency.

$$dS_t = \sigma\, dW_t\ , \tag{4.5}$$

$$\widehat{S}_t^{\nu_t} = S_t + \eta_t - k\,\nu_t\ , \tag{4.6}$$

$$d\eta_t = -\beta\,\eta_t\, dt + \sigma_\eta\, dW_t^\eta\ , \tag{4.7}$$

where $d\left[W, W^\eta\right]_t = \rho\, dt$ is the relation between the underlying price fluctuation and the latency impact is ρ, which is assumed to be constant.

Here, the underlying price follows a standard Brownian motion with volatility σ, and the execution price will be influenced by two facts: the temporary impact caused by walking the limit order book and the latency. Here, we assume the temporary impact only depends on the investors' trading speed. The parameter k is constant and fixed in our model.

The latency impacts we sustain depend on the market conditions mentioned previously, and the impacts can be either negative or positive. In the latency model, the mean is zero, and β is the speed of the mean reversion, while σ_η is the volatility of the latency impact.

We use the parameter set $\psi = 0.9$, $\theta = 0$, $\lambda = 1$ and $J_{max} = 10$ in the delay time model with reference to [4]. The reason why we set $\theta = 0$ is that we make the assumption that there exists an ideal situation where delay time can be dispelled. We then calibrate the parameters in our model accordingly and get the following results. The value in the brackets is the MSE of the estimation.

Ticker	AAPL	AMZN	FB	GOOG	INTC	MSFT	NFLX	NVDA	TSLA
$\widehat{\beta}$	1.0143	1.1100	1.1502	1.0727	0.1481	1.1664	1.0693	1.1174	1.0813
	(2.4e-5)	(0.0031)	(2.4e-5)	(0.0031)	(8.8e-6)	(1.2e-5)	(5.4e-5)	(4.4e-5)	(0.0012)
$\widehat{\sigma_\eta}$	0.0143	0.1575	0.0148	0.1616	0.0092	0.0108	0.0207	0.0187	0.1029
	(0.0028)	(0.0629)	(0.0029)	(0.0538)	(0.0011)	(0.0014)	(0.0072)	(0.0060)	(0.0276)

Table 4.3: Coefficients in the Stochastic Latency Impact Model: This is the case where $\psi = 0.9$, $\theta = 0$, $\lambda = 1$ and $J_{max} = 10$ in the delay time model. All the parameters are estimated by using the nine stocks in NASDAQ. All the data are taken from 12th April, 2017, in the HF form.

4.3.2 Performance criterion and value function.

Suppose we have a large quantity of $\mathfrak{N}$ shares to be liquidated. We denote the liquidation speed which is under the control of the investor, i.e. $\nu = (\nu_t)_{\{0\leq t\leq T\}}$, and denote the inventory by $Q^\nu = (Q^\nu_t)_{\{0\leq t\leq T\}}$. This is affected by how fast the investor trades, and satisfies

$$dQ^\nu_t = -\nu_t\, dt\,, \qquad\qquad Q^\nu_0 = \mathfrak{N}\,. \tag{4.8}$$

The investor's objective is to maximize the cash proceeds from selling shares. We denote the cash process by $X^\nu = (X^\nu_t)_{\{0\leq t\leq T\}}$. This satisfies the SDE

$$dX^\nu_t = \widehat{S}^\nu_t\, \nu_t\, dt\,, \qquad X^\nu_0 = x\,. \tag{4.9}$$

The investor's performance criterion is

$$H^\nu(t,x,S,\eta,q) = \mathbb{E}_{t,x,S,\eta,q}\left[X_T + Q^\nu_T\,(S^\nu_T - \alpha\, Q^\nu_T) - \phi\int_t^T (Q^\nu_u)^2\, du\right], \tag{4.10}$$

where the operator $\mathbb{E}_{t,x,S,\eta,q}[\,\cdot\,]$ represents the expectation when (with a slight abuse of notation) $X_t = x$, $S_t = S$, $\eta_t = \eta$ and $Q_t = q$. Here α is the terminal penalty for the final execution and ϕ is the inventory penalty. Its value function is

$$H(t,x,S,\eta,q) = \sup_{\nu\in\mathcal{A}} H^\nu(t,x,S,\eta,q)\,, \tag{4.11}$$

where $\mathcal{A}$ is the set of admissible strategies consisting of $\mathcal{F}$-predictable processes such that $\int_0^T |\nu_u|\, du < +\infty$, $\mathbb{P}-$a.s..

4.3.3 Dynamic Programming Equation

The dynamic programming equation that we want to solve is given by

$$\begin{aligned}\partial_t H + \frac{1}{2}\sigma^2 \partial_{SS} H + \frac{1}{2}\sigma_\eta^2 \partial_{\eta\eta} H + \rho\,\sigma\,\sigma_\eta \partial_{S\eta} H - \phi\, q^2 - \beta\,\eta\,\partial_\eta H \\ + \sup_{\nu\in\mathcal{A}} \{(S + \eta - k\,\nu)\,\nu\,\partial_\nu H - \nu\,\partial_q H\} \;\; = \;\; 0\;, \end{aligned} \tag{4.12}$$

with terminal condition $H(T, x, S, \eta, q) = X_T + Q_T\,(S_T - \alpha\, Q_T)$.

The first order condition gives us the initial optimal liquidation speed as

$$\nu^* = \frac{S+\eta}{2\,k} - \frac{\partial_q H}{2k\,\partial_x H}\;. \tag{4.13}$$

From looking at the terminal condition, and the way q enters into the DPE, we assume that the value function can be written as a quadratic function in q,

$$H(t, x, S, \eta, q) = x + q\,S + h_0(t,\eta) + h_1(t,\eta)\,q + h_2(t)\,q^2\;, \tag{4.14}$$

with the terminal condition $h_2(T) = -\alpha$ and $h_1(T,\eta) = h_0(T,\eta) = 0$.

Then the supremum part is rewritten as

$$\sup_{\nu\in\mathcal{A}} \{(S + \eta - k\,\nu)\,\nu\,\partial_\nu H - \nu\,\partial_q H\} = \frac{1}{4k}\,(\eta - h_1 - 2h_2 q)^2\;. \tag{4.15}$$

Since we set the ansatz in the form (4.14), the cross term of the second derivative of H is zero, i.e. $\partial_{S\eta} H = 0$. As a result, we do not need to account for the correlation between the underlying price and latency impact.

Moreover, upon applying the ansatz to the above non-linear PDE (4.12), we find

that

$$\begin{aligned} & q^2 \quad \left(\partial_t h_2 - \phi + \frac{1}{k} h_2^2\right) \\ + \quad & q \quad \left(\partial_t h_1 - \beta\,\eta\,\partial_\eta h_1 + \frac{1}{2}{\sigma_\eta}^2 \partial_{\eta\eta} h_1 + \frac{1}{k} h_1\, h_2 - \frac{1}{k}\eta\, h_2\right) \\ + \quad & \left[\partial_t h_0 - \beta\,\eta \partial_\eta h_0 + \frac{1}{2}{\sigma_\eta}^2 \partial_{\eta\eta} h_0 + \frac{1}{4k}\left(\eta - h_1\right)^2\right] = 0\ . \end{aligned} \tag{4.16}$$

4.3.4 Solving the PDE System

Since the equation in (4.16) must be valid for each $q > 0$, each term in brackets must be individually vanished. This provides us with the PDE system in h_0, h_1 and h_2 as follows:

$$\partial_t h_2 - \phi + \frac{1}{k} h_2^2 = 0, \qquad h_2(T) = -\alpha\ , \tag{4.17}$$

$$\partial_t h_1 - \beta\,\eta\,\partial_\eta h_1 + \frac{1}{2}{\sigma_\eta}^2 \partial_{\eta\eta} h_1 + \frac{1}{k} h_1\, h_2 - \frac{1}{k}\eta\, h_2 = 0, \qquad h_1(T,\eta) = 0\ , \tag{4.18}$$

$$\partial_t h_0 - \beta\,\eta \partial_\eta h_0 + \frac{1}{2}{\sigma_\eta}^2 \partial_{\eta\eta} h_0 + \frac{1}{4k}\left(\eta - h_1\right)^2 = 0, \qquad h_0(T,\eta) = 0\ . \tag{4.19}$$

The first equation in (4.17) is a Ricatti type equation. See details in Chapter 2. We rearrange the equation as follows.

$$\frac{\partial_t h_2}{\sqrt{k\phi} - h_2} + \frac{\partial_t h_2}{\sqrt{k\phi} + h_2} = 2\sqrt{\frac{\phi}{k}}\ , \tag{4.20}$$

Integrating both sides from t to T, with the terminal condition $h_2(T) = -\alpha$, we can find the solution.

We rewrite (4.18) into a linear equation in η, i.e. $h_1(t,\eta) = l_0(t) + l_1(t)\,\eta$, where $l_0(T) = l_1(T) = 0$. This gives

$$\eta\left(\partial_t l_1 - \beta\, l_1 + \frac{1}{k}\, l_1\, h_2 - \frac{1}{k} h_2\right) + \left(\partial_t l_0 + \frac{h2}{k} l_0\right) = 0\ . \tag{4.21}$$

We have two ODEs,

$$\begin{cases} \partial_t l_1 - \beta\, l_1 + \frac{1}{k}\, l_1\, h_2 - \frac{1}{k} h_2 = 0 \ , \\ \partial_t l_0 + \frac{h2}{k} l_0 = 0 \ . \end{cases} \tag{4.22}$$

Multiplying by the integral factor $e^{-(\beta - \frac{h_2}{k})\, t}$ gives

$$\int_t^T e^{-(\beta - \frac{h_2}{k})\, s}\, l_1(s) ds = \int_t^T e^{-(\beta - \frac{h_2}{k})\, s} \frac{h_2}{k} ds \ , \tag{4.23}$$

Then we obtain the solution $l_1(t) = -\int_t^T e^{-(\beta - \frac{h_2(s)}{k})(s-t)\ \frac{h_2(s)}{k}} ds$.

Since $l_0(T) = 0$, we see that $l_0(t) = 0$.

Finally, we rewrite (4.19) in the form $h_0(t,\eta) = m_0(t) + m_1(t)\,\eta + m_2(t)\,\eta^2$, where $m_0(T) = m_1(T) = m_2(T) = 0$. This gives

$$\eta^2 \left(\partial_t m_2 - 2\,\beta\, m_2(t) + \frac{1}{4k}\right) + \eta \left(\partial_t m_1 - \beta\, m_1(t) - \frac{1}{2k}\, h_1\right) + \left(\partial_t m_0 + \sigma^{\eta^2}\, m_2(t) + \frac{1}{4k} h_1^2\right) = 0 \ . \tag{4.24}$$

The solution of it is

$$m_2(t) = -\frac{1}{8\, k\beta} \left(e^{-2\beta(T-t)} - 1\right) \ , \tag{4.25}$$

$$m_1(t) = -\frac{1}{2k} \int_t^T e^{-\beta(s-t)\, h_1(s,\eta)}\, ds \ , \tag{4.26}$$

$$m_0(t) = \frac{1}{4k} \int_t^T \left(\sigma_\eta{}^2\, m_2(s) + \frac{1}{4k} h_1(s,\eta)^2\right) ds \ . \tag{4.27}$$

Based on the calculation mentioned above, we obtain the final solution of h_2, h_1 and h_0 as follows:

$$h_2(t) = \sqrt{k\,\phi}\, \frac{1 + \zeta\, e^{2\gamma(T-t)}}{1 - \zeta\, e^{2\gamma(T-t)}} \ , \tag{4.28}$$

$$h_1(t,\eta) = -\eta \int_t^T e^{-(\beta - \frac{h_2}{k})(s-t)}\, \frac{h_2}{k} ds \ , \tag{4.29}$$

$$h_0(t,\eta) \;=\; \frac{1}{4k}\int_t^T \left(\sigma_\eta{}^2\, m_2 + \frac{1}{4k}h_1^2\right) ds - \frac{\eta}{2k}\int_t^T e^{-\beta\,(s-t)\,h_1}\, ds - \eta^2\, m_2(t) \tag{4.30}$$

where $\gamma = \sqrt{\frac{\phi}{k}}$, $\zeta = \frac{\alpha+\sqrt{k\,\phi}}{\alpha-\sqrt{k\,\phi}}$ and $m_2(t) = -\frac{1}{8\,k\beta}\left(e^{-2\beta(T-t)} - 1\right)$.

Theorem 4.3.1 ***Verification.*** *Based on the value function in (4.11), we find the optimal trading speed is given by*

$$\nu_t^* = \frac{1}{2k}\left[\eta_{t^-}\left(1 - \int_t^T e^{-(\beta-\frac{h_2}{k})(s-t)}\,\frac{h_2}{k} ds\right) - 2\,h_2\,q\right] \;, \tag{4.31}$$

where $h_2(t) = \sqrt{k\,\phi}\,\frac{1+\zeta\,e^{2\gamma(T-t)}}{1-\zeta\,e^{2\gamma(T-t)}}$, $\gamma = \sqrt{\frac{\phi}{k}}$ *and* $\zeta = \frac{\alpha+\sqrt{k\,\phi}}{\alpha-\sqrt{k\,\phi}}$, *which is the optimal control we seek. Here,* η_{t^-} *is the latency impact we capture from the previous trading period.*

Proof. Since $x + q\,S + h_0(t,\eta) + h_1(t,\eta)\,q + h_2(t)\,q^2$ is clearly a classical solution, and the standard results imply that it suffices to confirm that this control is indeed an admissible strategy. Meanwhile, from the form of the optimal control in (4.13), we have

$$\nu^* = \frac{1}{2\,k}\left(\eta - h_1 - 2h_2\,q\right) \;, \tag{4.32}$$

and the explicit form of $h_2(t)$ and $h_1(t,\eta)$, thus we obtain the expansion form of ν^* above. □

4.4 Simulations of the Strategy with Stochastic Latency Impact

4.4.1 Optimal Strategy

In this section, we perform the simulation to show the behaviour of the optimal liquidation provided in (4.31). We assume that the trader will liquidate 1,000 shares of GOOG, i.e. $\mathfrak{N} = 1000$.

We use the GOOG data on 12th April, 2017, as an example. The coefficients we use here are the coefficients of the delay time model in the previous section, i.e. $\psi = 0.9$, $\theta = 0$, $\lambda = 1$ and $J_{max} = 10$. We also make some assumptions for the parameters as follows:

$$k = 0.001 \ , \quad \phi = 0.1 \ , \quad \alpha = 10$$

Here $T = 1$ means we focus on a single trading day, i.e. intraday trading. The other parameters such as $\beta = 1.0727$ and $\sigma_\eta = 0.1616$ can be found in Table 4.3.1.

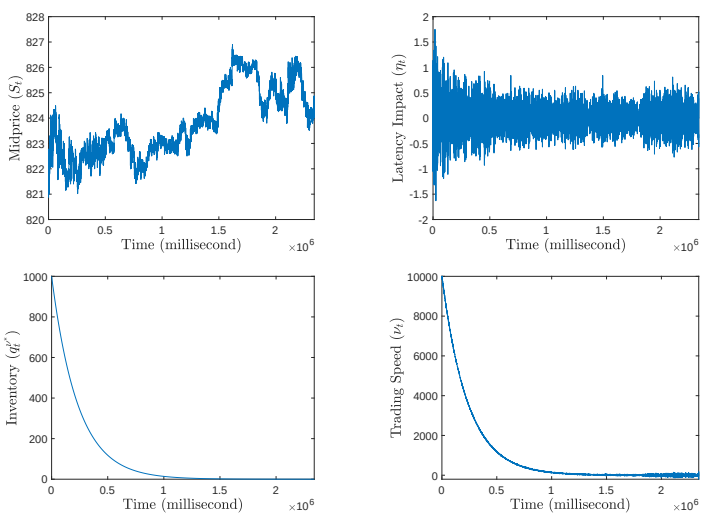

Figure 4.2: Sample path of underlying Stock Price, latency impact, Optimal Inventory and trading speed. The data is from GOOG on 12th April, 2017.

4.4.2 Simulations of Latency Impact

We run 1,000 times of the simulations of the latency impact and compare the performance with that in the Almgren-Chriss model. The optimal liquidation speed in the Almgren-Chriss model is

$$\nu_t^{AC} = \gamma \, \frac{\zeta \, e^{\gamma(T-t)} + e^{-\gamma(T-t)}}{\zeta \, e^{\gamma(T-t)} - e^{-\gamma(T-t)}} \, q_t^{\nu} \ , \tag{4.33}$$

where γ and ζ are defined as in the equation (4.28).

We use the following measurement to evaluate the performance

$$\frac{X_T^{\nu^*} - X_T^{AC}}{X_T^{AC}} \times 10^4 \ . \tag{4.34}$$

The unit here is basis points (bps).

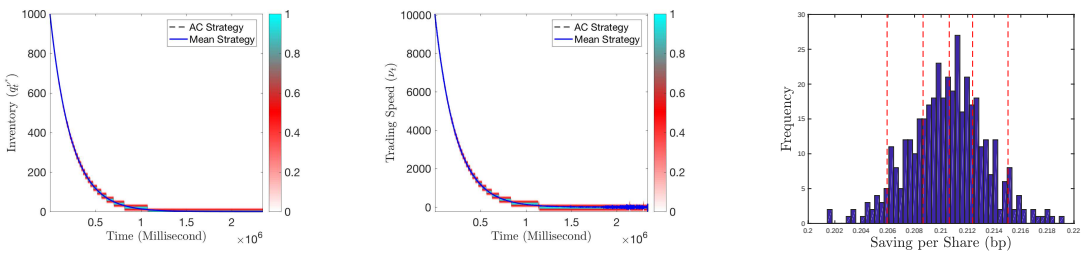

Figure 4.3: The simulation results and performance.

Figure 4.3 shows the simulation results. The first two are the heatmaps of the optimal inventory and the trading speed. From the figures, we see the strategy is very close to the Almgren-Chriss one, especially the mean of our strategy. The performance saving per share is always positive. The dashed red lines indicate the 5%, 25%, 50%, 75% and 95% quantiles moving from left to right.

4.4.3 Performance of the Strategy

Using the method elaborated in the last section, we show the performance of each stock as in Table . Here, μ_P and σ_P represent the mean and the standard deviation of the performance from 1000 simulated results, and the percentage represents the quantile of the performance.

Ticker	AAPL	AMZN	FB	GOOG	INTC	MSFT	NFLX	NVDA	TSLA
μ_P	0.2134	0.2180	0.2141	0.2120	0.2134	0.2148	0.2151	0.2101	0.2081
σ_P	0.0016	0.0030	0.0012	0.0026	0.0022	0.0016	0.0024	0.0031	0.0039
5%	0.2105	0.2134	0.2121	0.2074	0.2098	0.2122	0.2112	0.2049	0.2013
25%	0.2126	0.2161	0.2134	0.2103	0.2118	0.2136	0.2135	0.2079	0.2054
50%	0.2135	0.2178	0.2141	0.2124	0.2136	0.2148	0.2150	0.2100	0.2083
75%	0.2144	0.2199	0.2149	0.2138	0.2151	0.2157	0.2168	0.2123	0.2110
95%	0.2158	0.2236	0.2161	0.2158	0.2168	0.2180	0.2194	0.2155	0.2144

Table 4.4: Performance of the Strategy (in basis points).

4.5 Chapter Conclusions

We show the existence of latency impact by empirical evidence and employ an OU type process to model it, which follows the AR model in the discrete form. We then provide a closed form trading strategy to deal with this latency impact as it is encountered in the context of electronic trading system. The strategy discussed in this paper is established on the basis of general assumptions that are widely recognized in high frequency trading, for example, where MO are posted in every 10ms, while particularly, we add a stochastic delay time into the strategy to simulate the interval period caused by the travelling of eh signals.

For the delay time model, we discussed different levels of the lowest latency, which is the mean in the model. It is confirmed that there exists the autocorrelation in the latency impact, encouraging us to use the OU type process.

The optimal strategy is a linear form of latency impact. The improvement is not great, if it is at least statistically significant. However, if we focus on the market micro-structure and perform the simulation, we can detect the improvement based on the real data from nine stocks in nice NASDAQ stocks. The results show the improved performance when considering latency impact. To further improve the performance in future, we can set price that is not Brownian and execute a lager buy/sell orders.

Chapter 5

Pairs Trading of Cryptocurrencies

5.1 Introduction

A cryptocurrency is a digital asset designed to work as a medium of exchange, using cryptocurrency to secure the transactions and to control the creation of additional units of the currency [12]. Cryptocurrencies are also referred to as digital currency, token, cryptocoin, and e-money. Bitcoin was the first cryptocurrency traded in the market. It started to be traded in 2009 and its supply is limited to the total of 21,000,000 coins. Currently around 16,000,000 coins are in circulation.[1] The supply of Bitcoin increases when the activity that the so-called miners verify that the Bitcoin transactions are genuinely happens. Therefore, new Bitcoins are issued to pay for the mining service.

Nowadays, over 1,500 cryptocurrencies can be traded in different cryptocurrency exchanges that operate 7 days a week, 24 hours a day.[2] In traditional electronic foreign exchange (FX) markets and cryptocurrency exchanges, agents can buy or sell cryptocurrency pairs. Most cryptocurrency exchanges are electronic and the trade is done via a limit order book. The cryptocurrency pairs consist of either two cryptocoins or a cryptocoin and one of the major currencies, including USD, GPB,

[1] https://coinmarketcap.com/.
[2] Ibid.

Euros, JPY.

In this chapter we develop a 'Pairs Trading' strategy to take positions in two cryptocurrency pairs. As an application, we show the performance of the strategy for BTC/USD and ETH/USD. Here the quote currency for both pairs is the USD (US dollar), and the base currencies are Bitcoin (BTC) and Ethereum (ETH). Pairs Trading is a classic strategy which makes the most of the predictability of the joint, rather than the individual behaviour of the two financial instruments, which in this paper are cryptocurrency pairs. Pairs Trading algorithms profit from betting on the empirical fact that spread deviations tend to return to their historical or predictable level.

In our model, the spread between the two currency pairs is modelled by a co-integrating factor. However, finding pairs or collection of financial instruments that are co-integrated is not straightforward. In many cases a collection of assets, such as currencies and interest rates, may exhibit dynamics that are co-integrated, but within a short period of time the co-integration structure breaks down.

Co-integration, different from correlation, analyzes the movements in prices and identifies the degree to which two values are sensitive to the same mean or average price over a given time period. It doesn't indicate the direction that the pairs will move towards. Co-integration only measures whether or not the distance between them remains stable over time, i.e. if two cryptocurrency pairs are co-integrated then it is possible to form a stationary pair from some linear combination of crypto-pair A and crypto-pair B.

There are few of academic literatures on cryptocurrencies and their statistical properties. However, there are a large number of papers that probe into pairs trading strategies. Their contribution is either on the statistical aspects or the modelling of the relationship between pairs of assets, whilst a few of others look at how to dynamically take positions in the pairs. For example, [14] proposes a mean-reverting

Gaussian Markov chain model for the spread of a pair of assets. [13] puts forward a model for co-integrated asset prices and focuses on the valuation of options.

One of the previous papers which employs stochastic control techniques to trade pairs of co-integrated assets is that of [25]. The authors model the log-relationship between a pair of stock prices as an Ornstein-Uhlenbeck process and use it to formulate a portfolio stochastic control problem. More recently, [22] studies the optimal timing strategies for trading a mean-reverting price spread. The authors formulate an optimal double stopping problem to analyze the timing to start and subsequently liquidate the position subject to transaction costs. [21] analyzes a multiple entry-exit problem on a pair of co-integrated assets. The authors recast the sequence of optimal stopping problems as variational inequalities and performs extensive numerical simulations (as well as calibrate to a pair of dual-listed Chinese stocks) to illustrate how the optimal strategy behaves. The work of [26] considers two correlated assets whose spread is modelled by a mean-reverting process with stochastic volatility and show how the investor switches between holding no stocks, longing one stock shorting the other, and vice-versa.

[30] develops an optimal portfolio strategy to invest in two risky assets as well as the money market account, assume that log-prices are co-integrated, and find, in closed-form, the dynamic trading strategy maximizes the investor's expected utility of wealth. [25] models the log-relationship between a pair of stock prices as an Ornstein-Uhlenbeck (OU) process and solve a portfolio optimization based problem. The work of [17] analyses a cointegration-based statistical arbitrage model and applies the mean variance utility to measure the performance of their trading strategies. [31] establishes the portfolio which consists of a bank account and two co-integrated stocks and the objective is to maximize for a fixed time horizon, the expected terminal utility of wealth. Our paper is very close to this work, but with different model setup and underlying assets. [8] assumes that the drift of asset re-

turns consists of an idiosyncratic and a common drift component and generalise the model to allow the investor to trade in m co-integrated assets. [23] focuses on a multi-dimensional version of the model of [30]. Their underlying assets is the Bitcoins in the different exchanges, which are Bitstamp, BTC-e and itBit. However, what we will focus in this thesis is that we will only probe into the trading strategies of different cryptocurrency pairs in one exchange. In this Chapter, we continue to apply the methodology created by Cartea [4] to the two dimensions to solve the problem of cryptocurrencies.

The rest of the chapter is organised as follows: Section 2 presents the dynamics of the co-integrated assets and lists the empirical evidence of the co-integrated factor. Section 3 develops the investor's optimal control problem and derives in closed-form the optimal trading strategy. In Section 4, we employ high-frequency data of BTC and ETH to simulate the strategy and analyse the profit and loss (P&L) and conclude in Section 5.

5.2 Empirical Evidence

5.2.1 Co-integration test

Here, we use Johansen co-integration test to verify whether the co-integrated factors continuously exist between the underlying assets, [18] and [15]. Johansen test can be treated as a multivariate generalisation of the augmented Dickey-Fuller test. The generalisation is the examination of linear combinations of variables for unit roots, which can be used for the multi-variables. According to Johansen co-integration test, we suppose Π is the product of the vector of adjustment parameters β for the series itself and the vector of co-integrating vectors α as

$$\mathbf{\Pi} = \boldsymbol{\beta}' \, \boldsymbol{\alpha} \ , \tag{5.1}$$

If the matrix Π equals a matrix of zeroes, that is, $\Pi = 0$, then we can say that the variables are not co-integrated.

We use the matlab routine `jcitest`. We employ Apple, Inc. (AAPL) and Facebook, Inc. (FB) as the stock market's examples, and Bitcoin (BTC), Ethereum (ETH) and Litecoin (LTC) as the cryptocurrency market's examples. We run the Johansen test and get the results as follows. The data we use are all last trade price, picked in every 10 minutes for stocks and every 15 minutes for cryptocurrencies.

Table 5.1 shows the test statistics and their p-values in parenthesis. Here the null hypothesis is that the co-integrated factor is zero. If we set the confidence at 95%, then the stock pairs of AAPL& FB and AMZN & FB fail to reject the null hypothesis on 14th and 15th June, which means the co-integration is weak on these two days. And the cryptocurrency pair BTC & ETH shows co-integration with each other during this period.

Pairs/Date	12th June	13th June	14th June	15th June	16th June
AAPL & AMZN	20.9153	19.0932	11.0111	15.4948	11.3147
	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
AAPL & FB	9.8339	11.2344	2.9326	2.9120	10.0247
	(0.0023)	(0.0010)	(**0.0868**)	(**0.0880**)	(0.0021)
AMZN & FB	9.8867	14.9904	2.8654	3.3207	18.9225
	(0.0023)	(0.0010)	(**0.0905**)	(**0.0685**)	(0.0010)
BTC & ETH	4.6135	19.0237	10.9866	4.6651	5.1798
	(0.0318)	(0.0010)	(0.0010)	(0.0308)	(0.0230)
BTC & LTC	15.5892	22.0198	9.7561	5.7328	3.7312
	(0.0102)	(0.0010)	(0.0393)	(**0.2421**)	(**0.3242**)
ETH & LTC	17.2376	27.6714	15.6473	3.2198	19.7824
	(0.0113)	(0.0010)	(0.0010)	(0.8022)	(0.0233)

Table 5.1: Johansen test for AAPL & AMZN, AAPL & FB, AMZN & FB and BTC & ETH from 12th June to 16th June, 2017. Estimated p-values are shown in parenthesis.

Table 5.2 is the intraday Johansen test for these stocks and cryptocurrency pairs on 16th June. We took this date as an example, and the other days are similar. Compared with Table 5.1, although all of these pairs show that they are co-integrated, these relationship may change during that day.

Pairs/Time	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00	15:00-16:00
AAPL & AMZN	6.9004	8.8327	12.8229	7.3661	5.9426	3.6536
	(0.0089)	(0.0037)	(0.0010)	(0.0071)	(0.0148)	(**0.0560**)
AAPL & FB	5.1895	3.6299	12.4434	6.1947	5.8930	5.6696
	(0.0229)	(**0.0568**)	(0.0010)	(0.0131)	(0.0152)	(0.0174)
AMZN & FB	6.4322	4.4216	13.7305	14.5327	10.6357	7.0158
	(0.0114)	(0.0355)	(0.0010)	(0.0010)	(0.0013)	(0.0085)
Pairs/Time	0:00-4:00	4:00-8:00	8:00-12:00	12:00-16:00	16:00-20:00	20:00-24:00
BTC & ETH	7.1441	6.1554	6.0357	7.7783	18.1351	7.9289
	(0.0080)	(0.0133)	(0.0142)	(0.0054)	(0.0010)	(0.0049)
BTC & LTC	7.1441	6.1554	6.0357	7.7783	18.1351	7.9289
	(0.0080)	(0.0133)	(0.0142)	(0.0054)	(0.0010)	(0.0049)
ETH & LTC	7.1441	6.1554	6.0357	7.7783	18.1351	7.9289
	(0.0080)	(0.0133)	(0.0142)	(0.0054)	(0.0010)	(0.0049)

Table 5.2: Johansen test for AAPL & AMZN, AAPL & FB, AMZN & FB and BTC & ETH on 16th June, 2017

Here, we divide one trading day into 6 pieces. For the stock market is from 10am to 4pm, while the cryptocurrencies change in every four hours. We can find that AAPL & AMZN breaks the link from 3pm to 4pm and AAPL & FB fails to pass Johansen test from 11am to 12pm.

Table 5.3 shows the trading volume for the stocks and cryptocurrencies discussed in the Table 5.1 and 5.2. There are more than 500 cryptocurrencies' exchanges trading BTC and ETH, and Kraken is one of the top 10 exchanges all over the world. We can use market order (MO) or limit order book (LOB) to trade on Kraken.

Pairs/Date	12th June	13th June	14th June	15th June	16th June
AAPL	72,307,300	34,165,400	31,531,200	32,165,400	50,361,100
AMZN	9,447,200	4,580,000	3,974,900	5,373,900	11,472,700
FB	33,170,200	20,483,400	20,808,800	18,994,200	22,882,400
BTC	2,569,530,000	1,781,200,000	1,696,560,000	2,026,260,000	1,195,190,000
ETH	2,882,650,000	1,717,380,000	1,272,580,000	2,463,450,000	1,096,280,000

Table 5.3: Trading volume of AAPL, AMZN, FB, BTC and ETH from 12th June to 16th June, 2017 (USD)

Meanwhile, in the common exchange procedure of cryptocurrencies, we can figure out the exchange rate of BTC/USD, ETH/USD and ETH/BTC, with the last one offering us the foundation to obtain the co-integrated factor between the two underlying assets.

5.2.2 Relationship between BTC and ETH

Different from the stock market, we can find various pairs in the cryptocurrencies' market, such as BTC/USD, ETH/USD and ETH/BTC. By collecting the price data in every 15 minutes, Table 5.2 shows BTC and ETH price and their trend on 16th June, 2017.

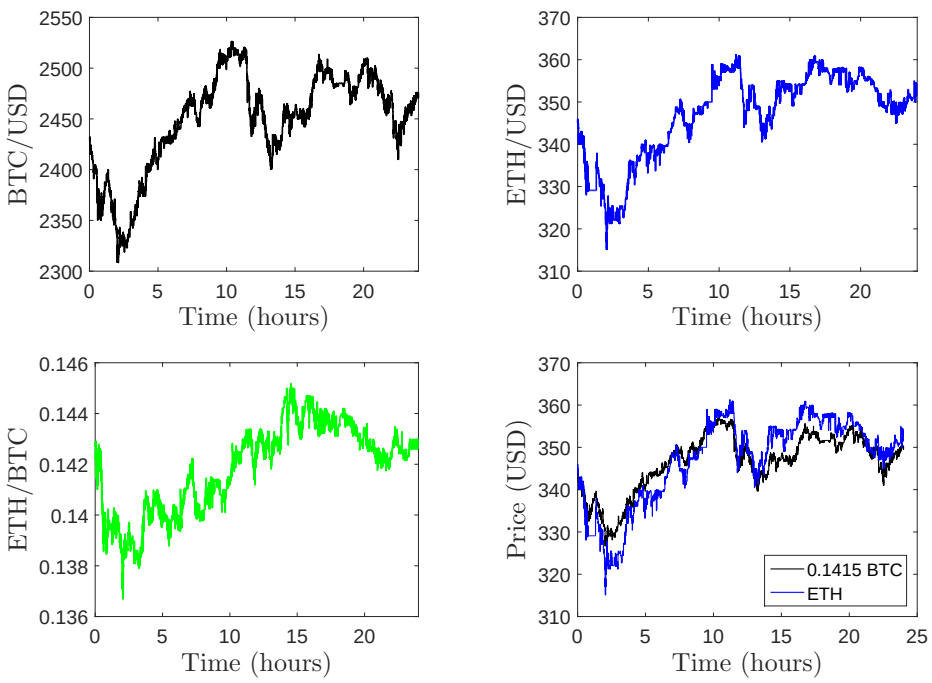

Figure 5.1: Date: 16th June, 2017. Top left: BTC/USD. Top right: ETH/USD. Bottom left: ETH/BTC. Bottom right: Scaled BTC (0.1415 BTC) and ETH.

We scale the BTC/USD to compare with ETH/USD in the bottom right of the figure. In order to get this scaled rate, we assume that the pair of ETH/BTC follows the AR(1) model, making X_t denote the ETH/BTC rate and X_t as follows:

$$X_{t+1} = \phi_0 + \phi_1 X_t + \epsilon \ , \tag{5.2}$$

where ϕ_0 and ϕ_1 is the parameter of lag one term. Here we use the matlab routine `arima` to estimate the parameter in the AR(1) model. By fitting the data of ETH/BTC on 15th June. We get the results as follows:

Parameter	Estimation	t Statistic	P-value
ϕ_0	0.000221422	4.02508	2.8575e-5
ϕ_1	0.998375	2477.46	1e-10

Table 5.4: Estimation of the AR(1) for ETH/BTC on 15th June

Based on 5.2, we can also change it into the continuous form

$$dX_t = \kappa_x \, (\theta_x - X_t) \, dt + \sigma_x \, dW_t \ , \tag{5.3}$$

where $\kappa_x = 1 - \phi_1 = 0.0016$, $\theta_x = \frac{\phi_0}{1-\phi_1} = 0.1363$ and σ_x is the standard derivation of the residual of regression in equation 5.2, which is 0.0253.

Since the mean reverting speed is very slow, we use the matlab routine `forecast` to forecast 'ETH/BTC' rate in every 15 minutes on 16th June. And then we take the average rate of these forecasting rates, 0.1415, as the scaled rate of BTC. Finally, we can reach to the difference between these co-movement pairs for the evidence of co-integrated factor.

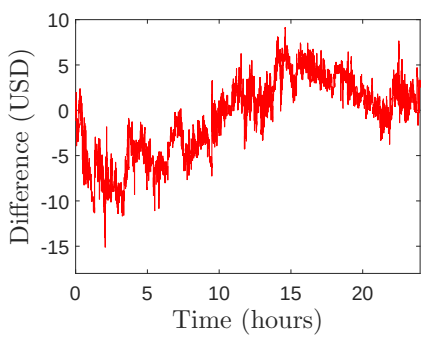

Figure 5.2: Difference between 0.1415 BTC and 1 ETH

In Figure 5.1, we already show the exchange rate of ETH/BTC, which means we can use this rate of BTC to exchange **ONE** ETH at that day that time. And we also have BTC and ETH in the same figure for comparison. Figure 5.2 shows the difference between them.

5.3 Co-integrated Log Prices with Short-Term Alpha

In this section, we explain the co-integrated factor attached in the pair of two underlying digital assets and find statistical evidence to confirm that the two assets are co-integrated with each other.

5.3.1 Co-integrated Factor

We assume the underlying cryptocurrencies S_1 and S_2 follow the SDEs whose drift of asset returns consists of a common component. One of them is the infrastructure token, i.e. BTC, etc. The other is ETH or any altcoin which can be exchanged by BTC, etc., via ICO. Thus, we have a pair of stochastic differential equations (SDEs) as follows:

$$\frac{dS_{1,t}}{S_{1,t}} = \alpha_t\, dt + \sigma_1\, dW_{1,t}\,, \tag{5.4}$$

$$\frac{dS_{2,t}}{S_{2,t}} = -\alpha_t\, dt + \sigma_2\, dW_{2,t} \tag{5.5}$$

and $(W_1, W_2) = (W_{1,t},\, W_{2,t})_{0\le t\le T}$ are standard correlated Brownian motions with instantaneous correlation ρ.

What needs to be noted here is that when $\alpha_t = 0$, both S_1 and S_2 are geometric Brownian motions with zero drift, hence they are martingales for $t \le T$. In general, however, α_t will be non-zero representing short-term deviations from martingale behaviour, and might be considered as a 'short-term alpha' affecting both underlying cryptocurrencies.

More specifically, we set this common component as below:

$$\alpha_t = a_1\, \log S_{1,t} + a_2\, \log S_{2,t}\,, \tag{5.6}$$

where a_1 and a_2 are constant, it becomes a linear relationship between the pair. Here, we derive the log price to satisfy the SDEs straightforward, i.e.

$$d\log S_{1,t} = \left(\alpha_t - \frac{1}{2}\,\sigma_1^2\right) dt + \sigma_1\, dW_{1,t}\,, \quad \text{and} \tag{5.7}$$

$$d\log S_{2,t} = \left(-\alpha_t - \frac{1}{2}\,\sigma_2^2\right) dt + \sigma_2\, dW_{2,t}\,. \tag{5.8}$$

Thus, the SDE for the short-term alpha can be found, applying Ito's lemma, i.e.

$$d\alpha_t = \kappa\,(\theta - \alpha_t)\, dt + \eta\, dW_t\,, \tag{5.9}$$

where

$$W_t = \frac{a_1\,\sigma_1}{\eta} W_{1,t} + \frac{a_2\,\sigma_2}{\eta} W_{2,t}\,,$$

is a standard Brownian motion and the constants

$$\kappa = a_2 - a_1, \qquad \theta = -\tfrac{1}{2}\frac{a_1\,\sigma_1^2 + a_2\,\sigma_2^2}{a_2 - a_1}, \qquad \eta = \sqrt{a_1^2\,\sigma_1^2 + 2a_1\,a_2\,\sigma_1\sigma_2\,\rho + a_2^2\,\sigma_2^2},$$

represent the mean-reversion rate, level of the short-term alpha process and the diffusion coefficient, respectively. We assume that $a_2 > a_1$ so that the process α_t is a mean-reverting process (as opposed to a mean-avoiding process).

5.4 Optimal Pairs Trading Problem

Once the pair of the underlying is found, we need to figure out a way to set the optimal trading rate for the pair. Assuming that there is no impact of the trading and we consider how to optimise the agent's utility of the expected wealth.

We use $m = (m_t)_{0\le t\le T}$ and $n = (n_t)_{0\le t\le T}$ to denote the inventory in the underlying S_1 and S_2 respectively, and denote $X = (X_t)_{0\le t\le T}$ to the cash process of an

agent's wealth which satisfies the SDE

$$dX_t^{m,n} = m_t\, dS_{1,t} + n_t\, dS_{2,t}\,.$$

Assuming that the trading is intraday, we set the risk free rate at zero.

The controlled system of SDEs is

$$\begin{aligned} dX_t^{m,n} &= \alpha_t(m_t S_{1,t} - n_t S_{2,t})dt + m_t\,\sigma_1\, S_{1,t} dW_{1,t} + n_t\,\sigma_2\, S_{2,t} dW_{2,t}\ , \\ d\log S_{1,t} &= \left(\alpha_t - \frac{1}{2}\,\sigma_1^2\right) dt + \sigma_1\, dW_{1,t}\,, \\ d\log S_{2,t} &= \left(-\alpha_t - \frac{1}{2}\,\sigma_2^2\right) dt + \sigma_2\, dW_{2,t}\,, \end{aligned}$$

with $d\,[W_{1,t}, W_{2,t}] = \rho\, dt$.

The agent will optimise his/her position in the assets directly, rather than the rate of trading, and has exponential utility function $u(x) = -e^{-\gamma x}$, for a constant coefficient of risk aversion γ. In our analysis we were not able to find a closed form solution for other types of utility functions. The agents' performance criterion is

$$H^{m,n}(t,x,y,z) = \mathbb{E}_{t,x,y,z}\left[-\exp\{-\gamma\, X_T^{m,n}\}\right]\,,$$

where $X_t = x$, $\log S_{1,t} = y$, $\log S_{2,t} = z$, $m_t = m$ and $n_t = n$. His/her value function is therefore

$$H(t,x,y,z) = \sup_{m,n\in\mathcal{A}} H^{m,n}(t,x,y,z)\,,$$

where the set of admissible strategies $\mathcal{A}$ contains strategies such that

$$\mathbb{E}\left[\int_0^T \left[(m_u\, S_{1,u})^2 + (n_u\, S_{2,u})^2\right]\, du\right] < \infty\,.$$

Alternatively, we can enforce the condition that m_t and n_t are $\mathbb{P}$-a.s. bounded.

5.4.1 The DPE and its solution

Employing the dynamic programming principle leads to the dynamic programming equation (DPE), and the value function should satisfy the following Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned} 0 \;=\; & H_t + \left(\alpha_t - \frac{1}{2}\,\sigma_1^2\right) H_y + \sigma_1\sigma_2\,\rho H_{yz} + \left(-\alpha_t - \frac{1}{2}\,\sigma_2^2\right) H_z + \frac{1}{2}\,\sigma_1^2\,H_{yy} + \frac{1}{2}\,\sigma_2^2\,H_{zz} \\ & + \sup_{m,n}\left\{ [\alpha_t(my - nz)]\,H_x + \frac{1}{2}\,\omega'\,\Sigma\,\omega\,H_{xx} + \omega'\Sigma_1 H_{xy} + \omega'\Sigma_2\,H_{xz}\}\right\}\;, \end{aligned} \tag{5.10}$$

subject to $H(T,x,y,z) = -e^{-\gamma\,x}$, where $H_{\bullet}$ and $H_{\bullet\bullet}$ mean the first and partial derivatives, respectively. And the matrix for the amount of each underlying asset is $\omega = \begin{bmatrix} my \\ nz \end{bmatrix}$ and the volatility matrices are $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\,\sigma_2\,\rho \\ \sigma_1\,\sigma_2\,\rho & \sigma_2^2 \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \sigma_1^2 \\ \sigma_1\,\sigma_2\,\rho \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} \sigma_2^2 \\ \sigma_1\,\sigma_2\,\rho \end{bmatrix}$.

Due to the presence of the co-integration factor, we expect that the value function depends on short term alpha instead of a combination of prices of both cryptocurrencies. Thus, we propose the trial solution $H(t,x,y,z) = -e^{-\gamma x}\,h(t,\alpha)$, where the state variable $\alpha_t = \alpha = a_1\,y + a_2\,z$. Then the DPE becomes

$$\begin{aligned} h_t = \sup_{m,n}\Bigg\{ & \left[\alpha\,(my - nz)\gamma\,h - \frac{1}{2}\gamma^2\left(m^2\,\sigma_1^2\,y^2 + 2\,m\,n\,\sigma_1\sigma_2\,yz\,\rho + n^2\,\sigma_2^2\,z^2\right)\right] h \\ & - \left[(a_1 - a_2)\,\alpha - \frac{1}{2}(a_1\,\sigma_1^2 + a_2\,\sigma_2^2) - (my\,\sigma_1^2 + nz\sigma_1\sigma_2\rho)\gamma\,a_1 - (nz\,\sigma_2^2 + my\,\sigma_1\sigma_2\,\rho)\gamma\,a_2\right] h_\alpha \\ & - \frac{1}{2}\left(\sigma_1^2\,a_1^2 + 2\,\sigma_1\sigma_2\rho\,a_1 a_2 + \sigma_2^2\,a_2^2\right)\,h_{\alpha\alpha}\Bigg\}\;, \end{aligned} \tag{5.11}$$

with the terminal condition $h(T,\alpha) = 1$.

It is straightforward to show that the initial optimal control (m^*, n^*) in the feed-

back form is

$$m^* = \frac{\alpha}{\gamma\,\sigma_1^2\,y\,(1-\rho^2)} + \frac{\alpha\,\rho}{\gamma\,\sigma_1\sigma_2\,y\,(1-\rho^2)} + \frac{a_1}{\gamma\,y}\,\frac{h_\alpha}{h}\ , \tag{5.12}$$

$$n^* = -\frac{\alpha}{\gamma\,\sigma_2^2\,z\,(1-\rho^2)} - \frac{\alpha\,\rho}{\gamma\,\sigma_1\sigma_2\,z\,(1-\rho^2)} + \frac{a_2}{\gamma\,y}\,\frac{h_\alpha}{h}\ . \tag{5.13}$$

Proposition 5.4.1 *Let the agent's value function satisfy equation (5.10). Then the optimal amount for each underlying is*

$$m_t^*\,S_{1,t} = \frac{\alpha_t}{\gamma\,\sigma_1^2\,(1-\rho^2)} + \frac{\alpha_t\,\rho}{\gamma\,\sigma_1\sigma_2\,(1-\rho^2)} + \frac{a_1}{\gamma}\,\frac{h_\alpha}{h}\ , \tag{5.14}$$

$$n_t^*\,S_{2,t} = -\frac{\alpha_t}{\gamma\,\sigma_2^2\,(1-\rho^2)} - \frac{\alpha_t\,\rho}{\gamma\,\sigma_1\sigma_2\,(1-\rho^2)} + \frac{a_2}{\gamma}\,\frac{h_\alpha}{h}\ . \tag{5.15}$$

Substitute this pair of amount into the HJB equation (5.11) above, the DPE reduces to the non-linear partial differential equation as follows:

$$\begin{aligned} h_t &= \frac{\alpha^2\rho}{\sigma_1\sigma_2\,(1-\rho^2)}h - \alpha\,(a_1-a_2)\,h_\alpha + \frac{1}{2}\sigma_1^2\,a_1\,h_\alpha + \frac{1}{2}\sigma_2^2\,a_2\,h_\alpha \\ &\quad + \frac{1}{2\,h}\left(\sigma_1^2\,a_1^2 + 2\,\sigma_1\sigma_2\rho\,a_1\,a_2 + \sigma_2^2\,a_2^2\right)\,h_\alpha^2 \\ &\quad - \frac{1}{2}\left(\sigma_1^2\,a_1^2 + 2\,\sigma_1\sigma_2\rho\,a_1 a_2 + \sigma_2^2\,a_2^2\right)\,h_{\alpha\alpha}\ . \end{aligned} \tag{5.16}$$

5.4.1.1 Solving the DPE

To solve this non-linear PDE, we set a new function $g(t,\alpha) = -\log h(t,\alpha)$, with the terminal condition $g(T,\alpha) = 0$. Then the equation (5.16) transforms into the linear PDE as follows:

$$\begin{aligned} \partial_t g &+ \left[\alpha\,(a_1-a_2) - \frac{1}{2}(\sigma_1^2\,a_1 + \sigma_2^2\,a_2)\right]\partial_\alpha g \\ &+ \frac{1}{2}\left(\sigma_1^2\,a_1^2 + 2\sigma_1\sigma_2\,\rho\,a_1\,a_2 + \sigma_2^2\,a_2^2\right)\partial_{\alpha\alpha} g + \frac{\alpha^2\,\rho}{\sigma_1\sigma_2\,(1-\rho^2)} = 0\ . \end{aligned} \tag{5.17}$$

Propose ansatz $g(t,\alpha) = k_2(t)\,\alpha^2 + k_1(t)\,\alpha + k_0(t)$; so that the equation (5.16)

becomes

$$\begin{aligned} 0 &= \left[\partial_t k_2 + 2k_2\,(a_1 - a_2) + \frac{\rho}{\sigma_1\,\sigma_2\,(1-\rho^2)}\right]\alpha^2 \\ &+ \left[\partial_t k_1 + k_1\,(a_1 - a_2) - k_2(\sigma_1^2\,a_1 + \sigma_2^2\,a_2)\right]\alpha \\ &+ \partial_t k_0 - \frac{k_1}{2}\left(\sigma_2^2 a_1 + \sigma_2^2 a_2\right) + k_2\left(\sigma_1^2\,a_1^2 + 2\sigma_1\,\sigma_2\,\rho\,a_1\,a_2 + \sigma_2^2\,a_2^2\right)\ , \quad (5.18) \end{aligned}$$

which leads to the following couple system of ODEs.

$$\begin{aligned} \partial_t k_2 + 2k_2\,(a_1 - a_2) + \frac{\rho}{\sigma_1\sigma_2\,(1-\rho^2)} &= 0 \quad (5.19) \\ \partial_t k_1 + k_1\,(a_1 - a_2) - k_2(\sigma_1^2\,a_1 + \sigma_2^2\,a_2) &= 0 \quad (5.20) \\ \partial_t k_0 - \frac{k_1}{2}\left(\sigma_2^2 a_1 + \sigma_2^2 a_2\right) + k_2\left(\sigma_1^2\,a_1^2 + 2\sigma_1\,\sigma_2\,\rho\,a_1\,a_2 + \sigma_2^2\,a_2^2\right) &= 0\ . \quad (5.21) \end{aligned}$$

Dealing with these ODEs with the final condition $k_2(T) = k_1(T) = k_0(T) = 0$, we have

$$\begin{aligned} k_2(t) &= \frac{\rho}{2\,(a_1 - a_2)\,\sigma_1\,\sigma_2\,(1-\rho^2)}\left[e^{2\,(a_1-a_2)\,(T-t)} - 1\right]\ , \quad (5.22) \\ k_1(t) &= \frac{(\sigma_1^2\,a_1 + \sigma_2^2\,a_2)\rho}{2\,(a_1 - a_2)^2\,\sigma_1\,\sigma_2\,(1-\rho^2)}\left[2\,e^{(a_1-a_2)\,(T-t)} - e^{2\,(a_1-a_2)\,(T-t)} - 1\right]\ , \quad (5.23) \\ k_0(t) &= \frac{(\sigma_1^2\,a_1 + \sigma_2^2\,a_2)^2\rho}{4\,(a_1 - a_2)^2\,\sigma_1\sigma_2\,(1-\rho^2)}\left\{\frac{2}{a_1 - a_2}\left[1 - e^{(a_1-a_2)\,(T-t)}\right]\right. \\ &- \left.\frac{1}{2\,(a_1 - a_2)}\left[1 - e^{2\,(a_1-a_2)\,(T-t)}\right] + (T-t)\right\} \\ &- \frac{(\sigma_1^2\,a_1^2 + 2\,\sigma_1\sigma_2\,\rho\,a_1\,a_2 + \sigma_2^2\,a_2^2)\,\rho}{2\,(a_1 - a_2)\,\sigma_1\sigma_2\,(1-\rho^2)}\left\{\frac{1}{2\,(a_1 - a_2)}\left[1 - e^{2\,(a_1-a_2)\,(T-t)}\right] + (T-t)\right\} \quad (5.24) \end{aligned}$$

Recalling $h(t,\alpha) = e^{-g(t,\alpha)}$, it is not difficult to find $h(t,\alpha) = e^{-\left[k_2(t)\,\alpha^2 + k_1(t)\,\alpha + k_0(t)\right]}$.

Theorem 5.4.2 ***Verification.*** *Based on the value function in (5.10), we find that the optimal trading amount for each underlying is given by*

$$m_t^* = \frac{1}{S_{1,t}}\left\{\frac{\alpha_t}{\gamma\,\sigma_1^2\,(1-\rho^2)} + \frac{\alpha_t\,\rho}{\gamma\,\sigma_1\,\sigma_2\,(1-\rho^2)} - \frac{1}{\gamma}\,a_1\,\left[2\,k_2(t)\,\alpha_t + k_1(t)\right]\right\}\ , \quad (5.25)$$

$$n_t^* = \frac{1}{S_{2,t}} \left\{ -\frac{\alpha_t}{\gamma\,\sigma_2^2\,(1-\rho^2)} - \frac{\alpha_t\,\rho}{\gamma\,\sigma_1\,\sigma_2\,(1-\rho^2)} - \frac{1}{\gamma}\,a_2\,\left[2\,k_2(t)\,\alpha_t + k_1(t)\right] \right\} \tag{5.26}$$

where $k_2(t)$ and $k_1(t)$ is defined as in equation (5.22) and (5.23).

Proof. Since $-e^{-\left\{\gamma\,x+\left[k_2(t)\,\alpha^2+k_1(t)\,\alpha+k_0(t)\right]\right\}}$ is clearly a explicit solution of equation (5.10), and the standard results imply that it suffices to check that this control is indeed an admissible strategy. Meanwhile, from the form of the optimal control in (5.12) and (5.13), we also get the explicit form of $\frac{h_\alpha}{h} = -\partial_\alpha g$, i.e.

$$\partial_\alpha g(t,\alpha) = 2\,k_2(t)\,\alpha + k_1(t)\ . \tag{5.27}$$

Thus, we obtain the expansion form of m_t^* and n_t^* above. □

In this section, we find the optimal amount of both cryptocurrencies. If we want to identify the amount of money invested in each underlying asset, we should multiple the price respectively, i.e. $(m_t^*\,S_{1,t})_{0\le t\le T}$ and $(n_t^*\,S_{2,t})_{0\le t\le T}$ are the optimal amount of cash flow into each cryptocurrency in the pair.

5.5 Simulations Performance of Strategy

5.5.1 Simulations of Co-integrated Factor

We have already showed the optimal strategy of the pairs traded in the last section. Here, we will consider a real case whose pair consisting of two cryptocurrencies– Bitcoin (BTC) and Ethereum (ETH).

We verify the co-integrated factor between these two cryptocurrencies in section 5.3, which is the in-sample path of the short term alpha. And we choose 'BTC/USD' and 'ETH/USD' on 16th June, 2017 to make the simulation and compare the results. By recalling the SDE in (5.9), we can simulate the co-integrated factor on this trading

day.

Assuming that we do the intraday trading, i.e. $T = 1$, the frequency of the execution is in every four seconds. As mentioned in the previous section, the exchanges of cryptocurrency are normally decentralized and open 7 days 24 hours. So, in our simulation the intraday trading period means 24 hours.

We set $a_2 = 1$, and find out other parameters in the model (5.6), (5.7) and (5.8) as follows by using ordinary least squares method.

Cryptocurrency	**a**	σ
BTC/USD	-0.1415	0.1293
ETH/USD	1	0.1872

Table 5.5: Parameters in the SDE of co-integrated factor.

Before running the simulation, we suppose that the coefficient of risk aversion is $\gamma = 0.5$. The results for other values of γ can be found in Table 5.6. Moreover, since $\kappa > 0$ the the process α_t is indeed mean-reverting. We also assume that the agent begins the day with exactly **ONE** Dollar, \$ 1(beginning of the day).

5.5.2 Performance

We employ cryptocurrency price data on 16th June, 2017 to analyse the performance of the trading strategy. We substitute the parameters in the last subsection and run 10,000 simulation of co-integrated factor α_t to find the amount of each cryptocurrency and to record the cash process of each run.

Here, we presume that the trading frequency is in every 4 seconds, which means we can trade 21,600 times every day. In Figure 5.3, the top left is one of the 10,000 simulation paths of short term alpha. The top right is the paired optimal amount of each cryptocurrency, m_t for BTC and n_t for ETH. This paired value is based on one simulation of the α_t.

The bottom left one is the cash process of wealth in USD unit. We made an assumption that our initial cash is \$1 in last section and the price value we used is the real price of BTC and ETH on 16th June, 2017, as shown in Figure 5.1. This cash flow path is one of the simulation paths as well. Among all these paths, we find that the terminal cash values are either positive or negative. Then we check the performance of 'ETH/BTC' in the bottom right. In this figure, we show the histogram of P&L of all these simulations. We also use the red dash line to show the mean of the performance, i.e. μ_R mean of all return. Here, $\mu_R = 0.4846$ USD. Since the initial cash is \$1, it is equivalent to the percentage, which means the average return rate is 48.46%. And the yellow dash line stands for the return of ETH/BTC on 16th June, 2017, i.e. $r_{ETH/BTC} = 0.002039$. ETH/BTC is exchanging BTC to ETH, i.e. long ETH and short BTC. This means that by following a buy-and-hold strategy on ETH/BTC from the beginning to the end of the day, you will have 0.2039% return.

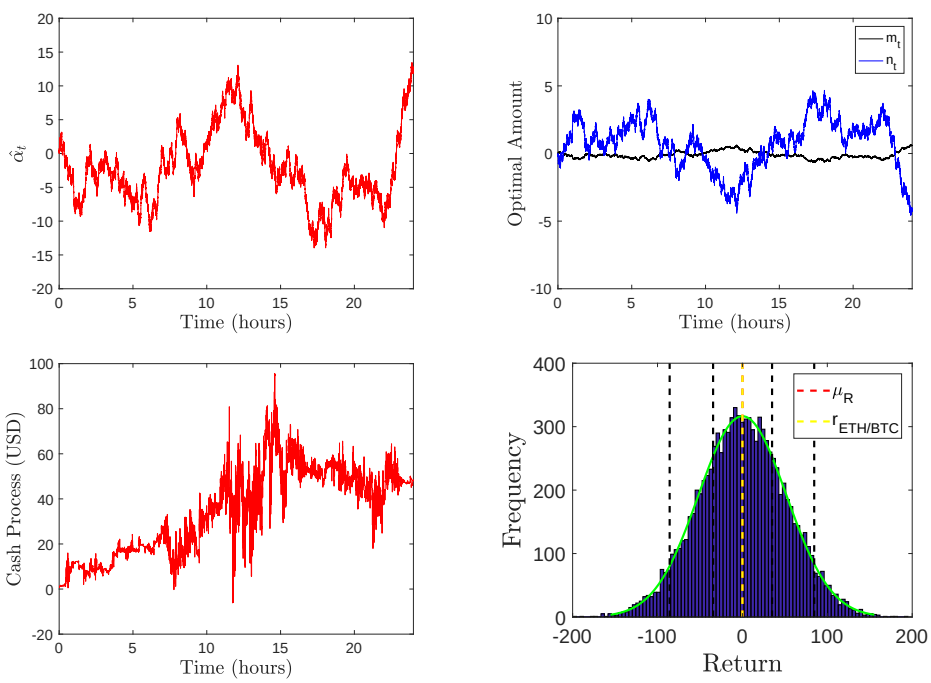

Figure 5.3: Top left: a simulation path of α_t on 16th June, 2017. Top right: the optimal amount of each cryptocurrencies in one simulation. Bottom left: the agent's cash process based on the path of short term alpha. Bottom right: the histogram of the profit & loss over 1,000 runs with normal distribution fitting.

5.5.3 Different Level of Risk Aversion Rate

In the last subsection, it is not difficult to find that the average performance is better than the one that only takes the long position in ETH/BTC. Here, we illustrate how the different levels of the risk aversion rate γ affects the performance of the trading.

γ	0.1	0.3	0.5	0.9
μ_R	2.4231	1.4541	0.4846	0.2692
σ_R	255.6527	153.3951	51.1305	28.4059
5%	-420.9126	-246.4685	-83.0892	-46.7405
25%	-169.5613	-100.12491	-33.2124	-18.7255
50%	4.2621	2.0874	0.9262	0.2744
75%	176.1892	104.8950	34.0796	19.7511
95%	412.2887	249.7608	84.2535	46.5861

Table 5.6: The performance based on the different γ.

If we calculate the performance based on the same simulation paths of the co-integrated factor with different risk aversion rate, we can find that the average performance μ_R decreases linearly when the γ increases. The variance of the performance can be found in the Table 5.6. It also shows the decay trend when γ becomes larger. This result is consistent with the principle of the investment. To put it in another way, the higher risk aversion is, the lower expected return can be.

5.6 Chapter Conclusion

We propose a model for the paired cryptocurrencies trading in this chapter and lay out the evidence that the stocks are not the good underlying asset to run the pairs trading strategy. The reason we choose the cryptocurrency is that it provides a perfect pair via the co-integrated factor between the icos' tokens.

In our model, we take Ethereum (ETH) as an example, to show the exchange between ETH and BTC. We establish the paired underlying asset price model for

these two cryptocurrencies. We use the empirical evidence to confirm that they are the good pair and there exists the co-integrated factor between the two. This short term alpha is in an OU type as assumed.

Finally, we use the historical data to estimate the parameters in the model and simulate 10,000 times of the co-integrated factor. Based on these simulations, we find out the optimal amount and the cash process. We compare the simulation performance with the return of simple long ETH short BTC, i.e. 'ETH/BTC', and confirm that the expected performance is better than direct trading 'ETH/BTC'. We also discuss the impact of choosing different risk aversion rate on the trading performance.

Chapter 6

Conclusions and Future Work

6.1 Final Conclusions

Trading impacts is the main topic of this thesis. We first go through the classic model, Almgren Chriss Model, for the benchmark. This model gives us the hint to upgrade the constant impact into a stochastic one. Then we introduce three problems that agents or traders will encounter in the process of trading. Here, we only consider that we use the MOs for the execution instead of LOBs, ignoring the price execution risk. And we use the HJB equation to solve the maximum utility problem in order to obtain the optimal control factors.

Optimal Execution with Stochastic Price Impact

In Chapter 3, we introduce the problem of the stochastic price impact. In the AC model, the price impact is constant. Here, we distinguish the temporary price impacts from the permanent ones, assume the linear correlation between these two impacts and use MLE to estimate the parameters for the stochastic model of price impacts. Based on the dynamic programming principle for the maximum utility function, we find out numerically the optimal trading strategies for the agents who want to liquidate a large number of shares without subject to significant price impact of large orders in a certain duration.

Optimal Execution with Stochastic Latency Impact

Chapter 4 shows the problem of the trading with stochastic latency impact. This impact is very common which means the delay time between the submission of the signal and the receipt of the order execution. We quantify the expectation of the difference between the submission price and the execution price, model this latency into the OU type process and then find the closed form solution to this optimal execution problem. For both problems, we upgrade the constant impacts into the stochastic form and compare the two results with the benchmark strategy. We find that if we model the impacts into a stochastic form, both will be more effective to help us to obtain better performing trading strategies.

Optimal Pairs Trading with Co-integrated impact for the cryptocurrencies

Chapter 5 presents a newly-emerging market of cryptocurrencies. Due to the nature of cryptocurrencies, new cryptocurrencies are introduced into the market with a pairing price with the existed mainstream cryptocurrencies, including Bitcoin (BTC) and Ethereum (ETH). As a result, being distinct from the traditional stock market where company shares in the same industry are expected to be correlated with each other and exposed to macro-economics and industry-related factors, cryptocurrencies are considered as co-integrated, thanks to the pairing price of their initial offering. Therefore, this study emphasizes on the myth, and the result is surprising. According to the result from the stochastic model of the short-term alpha and the co-integrated factor, we find an optimal closed form amount for both pairs of 'BTC/USD' and 'ETH/USD'.

6.2 Future Work

The thesis takes mathematical as well as empirical sides into consideration. For the future work, there are a number of directions that may lead the correspondent

researches. We will lay out some of the possibilities here.

- When we consider the optimal trading with the stochastic price impacts, we only use MOs in the thesis. In the real-world case, however, we also need to think about whether there is a dark pool for the agents to put their larger orders in or not. If the problem is faced with both lit and dark market, we need to take into account how deep the dark market is and how much we should post in the lit pool with the optimal trading speed.

- As illustrated in Chapter 3, we assume that the relationship between temporary and permanent impact is linear. Moreover, we can also try some non-linear models to link each other. This non-linear form may influence the entire execution strategy and further researches might be done on figuring out whether the influence is positive or negative, or even deeper. Our expectation is that the non-linear form will improve the trading performance since the assumption of non-linear relationship is more realistic. Or we model these two impacts respectively as two independent variables. We have only considered the cryptocurrencies as our underlying assets in the present study.

- In Chapter 5, what should be attached attention to is that combining cryptocurrencies and stocks in one model would be more practical because we have noticed the importance of hedging strategy through the process, which might instruct us to explore new trading strategies and derivative for cryptocurrencies in the future.

Chapter 7

Appendix

7.1 Numerical Explicit Scheme in Chapter 3

Here we assume $\boldsymbol{\mu} = 0$, which means the order flow is approximately symmetric for buy and sell side. So the net order flow is close to zero.

To solve the system of PIDEs, we first consider (3.16), and then (3.14) and (3.15) which will become simple ODEs if we get the value of (3.16). We introduce the explicit scheme for this finite difference. By using backward Euler for the time horizon and forward Euler for the temporary impact factor, we obtain

$$\frac{h_2(t,k) - h_2(t-\Delta t,k)}{\Delta t} + \beta\,(\xi - k)\,\frac{h_2(t,k+\Delta k) - h_2(t,k)}{\Delta k} - \phi + \lambda_k\,\mathbb{E}\,[h_2(t,k+\eta) - h_2(t,k)] + \frac{[(\ell_1 + \ell_2\,k) + 2\,h_2(t,k)]^2}{4k} \quad (7.1)$$

We employ backward Euler scheme for finite time horizon because we have the terminal condition for h_2, i.e. $h_2(T,k) = -\alpha$. We also require a boundary condition to employ an effective grid, thus we choose

$$\partial_k h_2(t,\bar{k}) = 0\,.$$

Here, $\bar{k}$ is the a maximum value for k, we discuss how to choose this $\bar{k}$ later. In this condition, when k is enormous, the value of h_2 will be independent of k. The backward finite difference form for this boundary condition is

$$h_2(t,\bar{k}) = h_2(t,\bar{k}-\Delta k)\,.$$

Let the domain of (t,k) be $[0,\,T]\times\left[\underline{k},\,\bar{k}\right]$ then we choose the minimum and maximum value of k as follows. From equation (3.5), the solution of k is

$$k_t = (k_0-\xi)\,e^{-\beta\,t} + \xi + \int_0^t \sigma_k\,e^{-\beta\,(t-s)}\,dW_s + \int_0^t \eta_{1+N_{s^-}}\,e^{-\beta\,(t-s)}d\widehat{N}_s\,. \tag{7.2}$$

When $t\to\infty$, the expected value and variance of k_t are

$$\begin{aligned}
\mathbb{E}\left[k_t\right] &= \xi + \frac{\lambda\,\mathbb{E}[\eta]}{\beta}\,, &(7.3)\\
\mathbb{V}\left[k_t\right] &= \mathbb{E}\left[\left(\int_0^t \sigma_k\,e^{-\beta_k\,(t-s)}\,dW_s\right)^2 + \left(\int_0^t \eta_{1+N_{s^-}}\,e^{-\beta_k\,(t-s)}d\widehat{N}_s\right)^2\right]\\
&= \frac{\sigma_k^2+\lambda\,\mathbb{E}[\eta^2]}{2\,\beta}\,. &(7.4)
\end{aligned}$$

When we run the numerical scheme, we need to set up a reasonable boundary for the impacts so as to contain as much possibility as possible. Hence, we fix the minimum value of k_t to half of the long term mean ξ, while the maximum value is fixed to ξ plus 3 times the standard deviation, i.e.

$$\begin{aligned}
\underline{k} &= \frac{\xi}{2}\,,\\
\bar{k} &= \xi + 3\sqrt{\frac{\sigma_k^2+\lambda\,\mathbb{E}[\eta^2]}{2\,\beta}}\,.
\end{aligned}$$

From the terminal condition, we can obtain $h_0(t,b,k) = h_1(t,b,k) = 0$ and $h_2(t,b,k)$. We get the surface of $g_2(t,b,k) = -\frac{(\ell_1+\ell_2\,k)+2\,h_2(t,k)}{2\,k}$ with different time

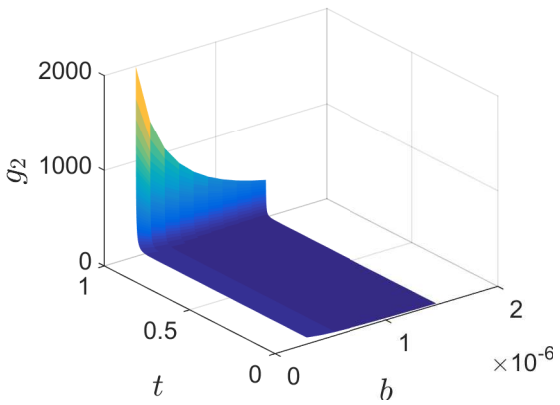

Figure 7.1: Surface of function g_2

and permanent impact in Figure 7.1. Here, we use the parameters of stock INTC which is estimated from the section 3.4.

7.2 Simualtion and Performance of other stocks in Chapter 3

The simulation and performance listed below is based on the common parameter set for the terminal execution penalty and inventory penalty, i.e. $\alpha = 10^3\,\xi$, $\phi = 10^3\,\xi$.

7.2.1 Optimal execution for FARO in Figure 7.2

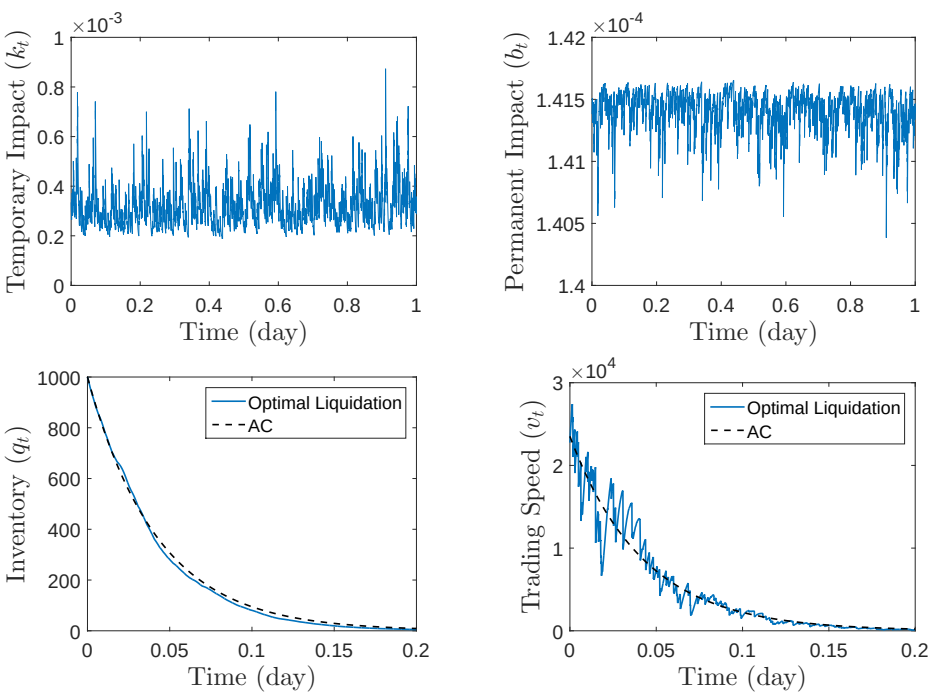

Figure 7.2: Optimal trading with stochastic temporary and permanent impact for FARO

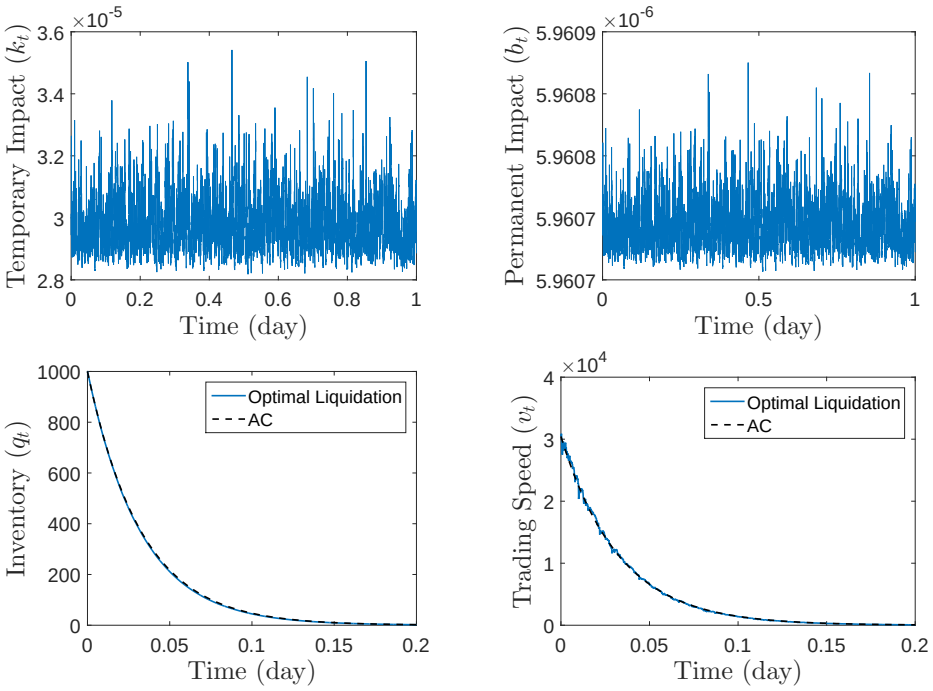

Figure 7.3: Optimal trading with stochastic temporary and permanent impact for NTAP

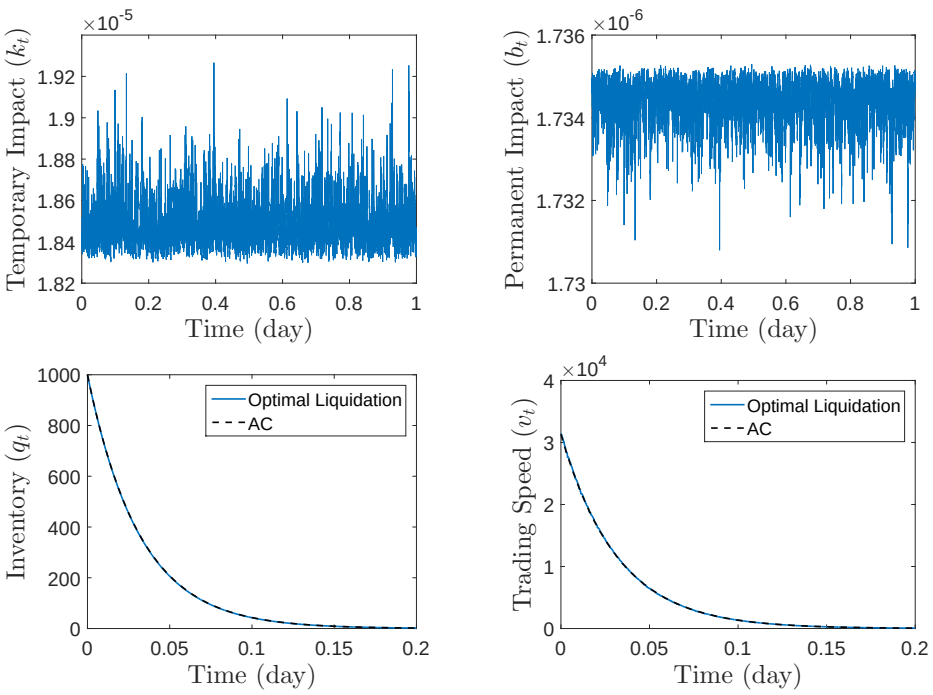

Figure 7.4: Optimal trading with stochastic temporary and permanent impact for ORCL

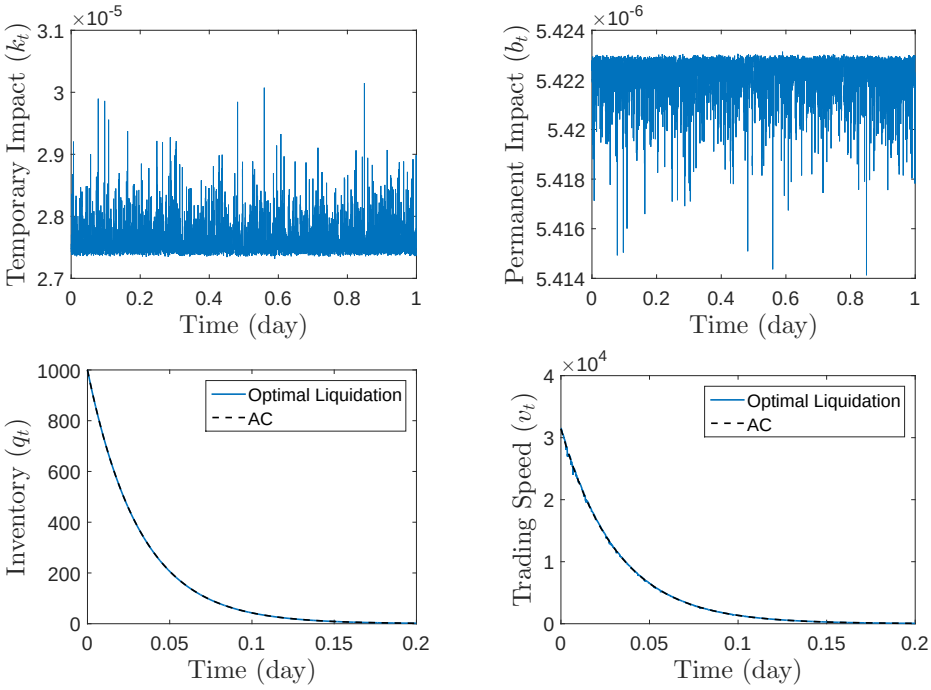

Figure 7.5: Optimal trading with stochastic temporary and permanent impact for SMH

7.2.2 Optimal execution for NTAP in Figure 7.3

7.2.3 Optimal execution for ORCL in Figure 7.4

7.2.4 Optimal execution for SMH in Figure 7.5

7.2.5 Performance of different stocks

Figure 7.6 shows the histogram of the savings per share in basis points for different stocks.

Table 7.1 shows the percentile value of each simulation performance for different stocks.

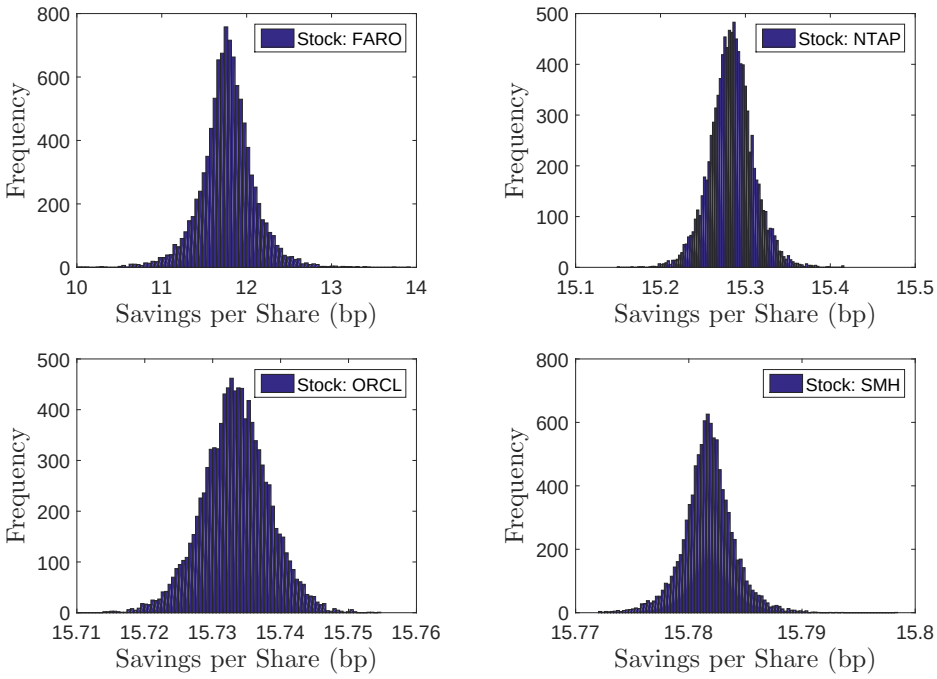

Figure 7.6: The savings per share measure in basis points for different stocks

	FARO	NTAP	ORCL	SMH
mean	11.7752	15.2839	15.7334	15.7817
stdev	0.3162	0.0259	0.0049	0.0023
5%	11.2683	15.2418	15.7252	15.7780
25%	11.6097	15.2681	15.7303	15.7805
50%	11.7736	15.2837	15.7334	15.7817
75%	11.9453	15.2991	15.7365	15.7830
95%	12.2880	15.3270	15.7414	15.7854
$X_T^{v^*} < X_T^C$	0%	0%	0%	0%

Table 7.1: Relative performance of the strategy in basis points for different stocks

7.3 Latency impacts for continuous Brownian Motion case in Chapter 4

Here, we simulate a Brownian Motion case to show the pattern of latency. Set $\sigma = 0.4$ for the Brownian Motion. We consider that the continuous case and this volatility value is the annualised volatility of the underlying asset.

We also assume that there are 250 trading days during a year and 6 trading hours per day. We post the market orders in every 20ms, and the delay time follows the model in equation (4.2). Keep the same parameter set with $\psi = 0.9$, $\theta = 10$, $\lambda = 1$ and $J_{max} = 20$.

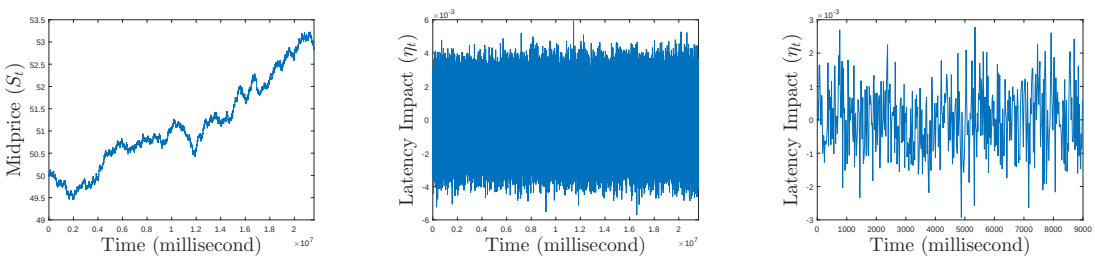

Figure 7.7: Continuous Case: the left one shows the underlying price. The middle one shows latency impacts during one trading day. The right one shows the latency impacts in the first 4000ms.

In Figure 7.7, the right plot looks very similar to an OU process. We find the result of ADF test also equals to 1, which means that we reject the null hypothesis. This tells us that latency impact is an auto correlated series for the continuous Brownian Motion case.

Bibliography

[1] Irene Aldridge. *High-frequency trading: a practical guide to algorithmic strategies and trading systems*, volume 459. John Wiley and Sons, 2009.

[2] Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[3] Robert Almgren, Chee Thum, Emmanuel Hauptmann and Hong Li. Direct estimation of equity market impact. *Risk*, 18(5752):10, 2005.

[4] Sebastian Jaimungal Álvaro Cartea and José Penalva. *Algorithmic and High-Frequency Trading*, volume 343. Cambridge University Press, 2015.

[5] Julius Bonart and Martin D Gould. Latency and liquidity provision in a limit order book. *Quantitative Finance*, pages 1–16, 2017.

[6] Álvaro Cartea, Ryan Donnelly and Sebastian Jaimungal. Algorithmic trading with model uncertainty. *SIAM Journal on Financial Mathematics*, 8(1):635–671, 2017.

[7] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *Available at SSRN 2557457*, 2015.

[8] Álvaro Cartea and Sebastian Jaimungal. Algorithmic trading of co-integrated assets. *International Journal of Theoretical and Applied Finance*, 19(06):1650038, 2016.

[9] Álvaro Cartea and Sebastian Jaimungal. A closed-form execution strategy to

target volume weighted average price. *SIAM Journal on Financial Mathematics*, 7(1):760–785, 2016.

[10] Álvaro Cartea and Sebastian Jaimungal. Incorporating order-flow into optimal execution. *Mathematics and Financial Economics*, 10(3):339–364, 2016.

[11] Ernie Chan. *Algorithmic trading: winning strategies and their rationale*. John Wiley & Sons, 2013.

[12] Usman W Chohan. Cryptocurrencies: A brief thematic review. 2017.

[13] Jin-Chuan Duan and Stanley R. Pliska. Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, 28(4):727 – 754, 2004.

[14] Robert J. Elliott, John Van Der Hoek and William P. Malcolm. Pairs trading. *Quantitative Finance*, 5(3):271–276, 2005.

[15] Robert F Engle and Clive WJ Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

[16] Xin Guo and Mihail Zervos. Optimal execution with multiplicative price impact. *SIAM Journal on Financial Mathematics*, 6(1):281–306, 2015.

[17] Markus Harlacher. Cointegration based statistical arbitrage. *Department of Mathematics, Swiss Federal Institute of Technology, Zurich, Switzerland*, 2012.

[18] Søren Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand, 1995.

[19] Barry Johnson. Algorithmic trading and dma, 2010.

[20] Charles-Albert Lehalle and Sophie Laruelle. *Market microstructure in practice*. World Scientific, 2013.

[21] Yaoting Lei and Jing Xu. Costly arbitrage through pairs trading. *Journal of Economic Dynamics and Control*, 56:1–19, 2015.

[22] Tim Leung and Xin Li. Optimal mean reversion trading with transaction costs and stop-loss exit. *International Journal of Theoretical and Applied Finance*, 18(03):1550020, 2015.

[23] Paul Sopher Lintilhac and Agnes Tourin. Model-based pairs trading in the bitcoin markets. *Quantitative Finance*, 17(5):703–716, 2017.

[24] Robert Merton. Optimum Consumption and Portfolio Rules in a Continuous Time Model. *Journal of Economic Theory*, 3:373–413, 1971.

[25] Supakorn Mudchanatongsuk, James A Primbs and Wilfred Wong. Optimal pairs trading: A stochastic control approach. In *American Control Conference, 2008*, pages 1035–1039. IEEE, 2008.

[26] Minh-Man Ngo and Huyen Pham. Optimal switching for pairs trading rule: a viscosity solutions approach. *Journal of Mathematical Analysis and Applications*, 441(1):403–425, 2016.

[27] Bernt Karsten Øksendal and Agnes Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, 2005.

[28] George Pruitt. *The Ultimate Algorithmic Trading System Toolbox*. Wiley, 2016.

[29] Ryan Riordan and Andreas Storkenmaier. Latency, liquidity and price discovery. *Journal of Financial Markets*, 15(4):416–437, 2012.

[30] Agnès Tourin and Raphael Yan. Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, 37(10):1972 – 1981, 2013.

[31] Agnès Tourin and Raphael Yan. Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, 37(10):1972–1981, 2013.

[32] Zhaodong Wang and Weian Zheng. *High-frequency trading and probability theory*. World Scientific, 2014.